

Tue Mar 19 12:36:56 1996

6

Tue Mar 19 12:36:56 1996

6

	Best Match	Local Similarity	24.4%	Pred. No.	8.41e-02;	Indels	Gaps	4,
Matches	29;	Conservative	33;	Mismatches	51;			
Db	255	dneirtlqsemagqdemiglnkdknleenkrtqgalaedkvnhlnkkaklestl	314					
Qy	270	EKVYSLIQNEVEKSVKSIOHNQCSFETTERQEMIRNKSILHQLRVIDQSQAERL	329					
Db	315	demeenlardeggirkvekrkrkkg-dlkata-etylrd-dlervkrd---leeglrre	367					
Qy	330	KELDKERIPFRONWEADSKMSKSVESLQNRYTELESUDKSAQVARNGLIGEQLSRRD	388					

Query Match		Score 3.2†;	Score 132;	DB 7;	length 527;
Best Local Similarity	24.4‡;	Pred. No.	8.41e-02;		
Matches	29;	Conservative	33;	Mismatches	51;
ID	R14183	standard; Protein:	527 AA.		
AC	R14183;				
DT	03-JAN-1992	(first entry)			
DE	Product of clone IRV-5.				
KW	EPIKVALIONNEURKOKSILOHNOICCSFTEIEKOMRINNEESKILHQLRVIDSREKL				329
Qy	dineirtlqlsqemaqqdemiglhnkdknleenkrtqeaedkvhnlnkkaklestl				314
Db	315 demeelnlardeckirkdvekkrkleg-dlkatq-etylvd-clervkrd---jeelqrke				367
Qy	330 KELDKIRPRQRQNWEADSKSSVSVSLQRNTELESUDKSAGQVARNTGLIESQLSRHD				388
RESULT	2				
ID	R14183	standard; Protein:	527 AA.		
AC	R14183;				
DT	03-JAN-1992	(first entry)			
DE	Product of clone IRV-5.				
KW	EPIKVALIONNEURKOKSILOHNOICCSFTEIEKOMRINNEESKILHQLRVIDSREKL				329
OS	Schistosoma mansoni.				
SH	immunogen; anti-IRV; fusion.				
FT	Domain	121..572			
FT	FT	alpha helix			
FT	FT	/label= N-glycos site			
FT	FT	Modified -site 218			
FT	FT	/label= N-glycos site			
PN	US5051254-A.				
PD	24-SEP-1991.				
PF	PR	30-SEP-1988; US-252075.			
PA	PA	(UYJO) JOHNS HOPKINS UNIV.			
PI	PI	Strand M;			
DR	DR	WPI: 91-30300341.			
PS	N-PSB2; QL13982	Polypeptide(s) specific for epitope(s) of Schistosoma mansoni –			
PT	PS	used as vaccines against bilharziasis.			
CC	CC	Claim 5; Fig 1B, 10pp, English.			
CC	CC	The clone was obtcd. by screening an adult worm cDNA library in			
CC	CC	lambda-gtl1 with anti-IRV sera raised against a subset of adult worm			
CC	CC	antigens. The sequence is thought to at least partially encode a			
CC	CC	schistosome myosin. It has 48% identity with the myosin heavy			
CC	CC	chain of C. elegans. The DNA can be used to produce a fusion protein			
SQ	SQ	with beta galactosidase sequences. The product can be used as an			
	Sequence	immunogen to produce vaccines against bilharziasis.			
	527 AA;	Sequence			

PR	07-JUL-1987; AD-002940.
PA	(Biotech Austr Pt, (CSIR) Commonwealth Sci and Ind Res Orgn,
PI	Cobon GS, Austin RA, O'Donnell IJ, Frenkel MJ, Kennedy WRK, Savin KW,
PI	Wagland BM;
DR	WPT; 89-039628-05.
DR	N-PSDB; N91211.
PT	protein derived from parasitic nematode species - used to provide
PT	protective immunity against nematode parasites of man and animals.
PS	Claim 4; Page 39; 57pp; English.
CC	Recombinant organisms, contg. DNA encoding the protein can be used to
CC	produce polypeptide which is capable of protecting guinea pigs
CC	against parasitism from <i>T. colubriformis</i> .
CC	See also P94260, and P94368-P94369.
SQ	Sequence 320 AA;
RESULT	3.1%; Score 128; DB 1; Length 320;
ID	Best Local Similarity 17.1%; Pred. No. 1.63e-01;
Matches	Mismatches 35; Indels 3; Gaps
18; Conservative	
Db	11 mkimekdahaldaradaeskvrqtkiteklerveelrdtqkkmmqtenddkdglasatsq 70
:: : :	
Qy	299 IEIERQKEMLRN-ESKILHQVRIDSQAEKILEDEIRIPRQRNWEA-DSMKSSVES 355
Db	71 lieekkvqgeeaevahnrrmtileeleraeerlkiateklee 115
: :	
Qy	356 LQNRVTELESVDKSAGOVARNITGLIESLQLSRDHOMLSVHDIRLAD 400
OS	
FH	R66930; standard; Protein; 885 AA.
Key	01-SEP-1995 (first entry)
DE	AMML chromosome inv(16) product.
KW	AMML; acute myelomonocytic leukemia; chromosome-16; inversion; inv(16); CBF-beta; CBF gene; transcription factor; myosin; MYH11; SMHC.
KW	Homo sapiens.
FT	Location/Qualifiers
Peptide	1..164
/label= CBFB	
FT	
Peptide	165..885
/label= MYH11	
FT	
W02004067-A.	
PD	09-FEB-1995.
PF	26-JUL-1994; US089869.
PR	29-JUL-1993; US089869.
PA	(UNM) UNIV MICHIGAN.
PA	(UTEX) UNIV TEXAS SYSTEM.
PI	Clyinton D, Collins FS, Liu P, Siciliano MJ;
DR	WPT; 95-082178/11.
DR	WPT; 084389.
PT	Novel DNA spanning the pericentromeric inversion of chromosome 16 -
PT	for the screening of acute myeloid leukaemia
PS	Claim 4; Page 34-38; 78pp; English.
CC	PCR was performed on total cellular RNA from 5 AMML patients having
CC	a pericentromeric inversion of chromosome-16, M4eo subtype. Sequencing
CC	showed the inv(16) fusion to comprise a sequence from the CBF
CC	gene, encoding a novel transcription factor, and the MYH11 gene,
CC	encoding smooth muscle myosin heavy chain. In 1 patient, nt 1-492
CC	of the CBF gene were fused to nt 994 of MYH11 (shown in

Tue Mar 19 12:36:56 1996

2

Tue Mar 19 12:36:56 1996

KW Malaria; sporozoite; vaccine; exoerythrocytic parasites;
KW tetanus toxoid.
OS Plasmodium falciparum.
PN W0906130-A.
PD 14-JUN-1990.
PF 30-NOV-1991 005335.
PR 12-APR-1989; US-337204.
PA (BIO) Biomedical Res. Inst.
PI Hollindale MR;
DR WPI; 90-203624927.
DR N-PSDB; Q05140.
PT Novel malaria sporozoite antigenic protein -
PT useful as vaccine against sporozoite(s) and exoerythrocytic
PT parasites.
PS Disclosure; p; English.
CC Antigen, preferably linked at the C-terminal to a carrier such as
CC tetanus toxoid, may be used as a vaccine against the malarial
CC sporozoite.
SQ Sequence 462 AA;

Query Match 2.7%; Score 114; DB 1; Length 462;
Best Local Similarity 21.9%; **Pred.** No. 1.57e+00;
Matches 32; **Conservative** 46; **Mismatches** 60; **Indels** 8; **Gaps** 8;

Db 204 lakekliqeqgsdileerrakelklqeqg-q-dleedrakelklqeqgsdileedrrakelklq 261
Qy 269 LER-KVSLIONESVEKNSKIQSLHNLQICSEIEER-OKEMRNNESKILHLQRVIDSA 326

Db 262 eqsksaleqrakekliqeqgsdileertkastetlqeragdqleerrakelklqeqgsdile-q 320
Qy 327 ERLKEELDE-I-RP-FRQNNEWEADSMKSSQVSLEONRVTELESVDKSAGQAVARTGLIESQ 383

Db 321 errakekliqeqgrdleqerlakerlakelq 346
Qy 384 LSRHDQMSVHDIRLADMIGFQVILE 409

RESULT 13
ID P90954 standard; protein; 769 AA.
AC P90954;
DT 27-Feb-1990 (first entry)
DE Yeast topoisomerase I cDNA
KW Sclerodermata.
OS Saccharomyces cerevisiae
PN W0809222-A.
PD 05-OCT-1989.
PF 22-MAR-1989 001116.
PR 23-MAR-1988; US-172159.
PA (BRIG) Brigham and Women's Hospital; (UYJO) John's Hopkins Univ.
PI Farshaw MC; D'Arpa P;
DR WPI; 89-30350042.
PT Cloned cDNA encoding eukaryotic topoisomerase I - useful for large scale
PT prodn. by recombinant methods
PS Disclosure; fig. 6; 28pp; English.
CC The cDNA of this can be spliced into DNA vectors and used to transform
CC hosts for high yield. This has high homology with human topoisomerase I.
SQ Sequence 769 AA;

Query Match 2.7%; **Score** 111; **DB** 1; **Length** 769;

Best Local Similarity 24.7%; **Pred.** No. 2.52e+00;
Matches 23; **Conservative** 26; **Mismatches** 41; **Indels** 3; **Gaps** 3;
Db 644 feke-ellpesqlkew-lekvdekkqefekelktgevelksvnsvekikavkleqr 701
Qy 301 IERQKMLRNNESKILHLQRVIDSAEKKELKEIREFRQNWEEDSMKSSQVESIQRV 360

Db 702 -qtssiqkdkkeensqvsigtkskinyidprlv 733
Qy 361 TELESVDKSAGOVARNGLLESQISRHDDOMLSV 393

RESULT 14
ID R30729 standard; Protein; 870 AA.
AC R30729;
DT 20-MAY-1993 (first entry)
DE p100 protein from human herpes virus type 6.
KW antibody; monoclonal antibody; ELISA assays; CMV; cytomegalovirus.
OS Human herpes virus type 6.
PN EP-52421-A.
PD 27-JAN-1993.
PF 15-JUN-1992; 110047.
PR 08-JUL-1991; EP-111338.
PA (BRHW) BEHRINGERWERKE AG.
PI Fleckenstein B, Neipel F;
DR WPI; 93-028331/04.
PT Human herpes virus type 6 protein p100 DNA sequence - useful in prophylaxis, treatment and differential diagnosis of human herpes virus-6 infections.
PT Claim 1; Page 12; 25pp; English.
CC This sequence is the p100 protein from human herpes virus type 6. The protein and antibodies to it can be used for treatment or prevention of HHV-6 infections. The DNA, protein and Ab are also useful in eg. ELISA assays esp. for differentiating between HHV-6 CC and cytomegalovirus infections. These assays are more sensitive CC and specific than immunofluorescence methods currently used.
SQ Sequence 870 AA;

Query Match 2.7%; **Score** 111; **DB** 6; **Length** 870;
Best Local Similarity 26.4%; **Pred.** No. 2.52e+00;
Matches 32; **Conservative** 34; **Mismatches** 47; **Indels** 8; **Gaps** 7;

Db 678 qrvnnilsqftnlisqntslrgeqkqiedlqn-grtqelklatenkagrseseanvekil-evsn 735
Qy 319 QRVID--SQAEKIEL-DRKTRPQQNWEADSMKSSQVSLEONRVTELESVDKSAGOVAR 375

Db 736 pgdmfn-friqndlsvqpfripadsls-reldsafkaldlkpgngereidle 793
Qy 376 NTGLESQSLRSRDQMLSVHDIRLADMIGFQVILETASYNGVLWIKIRDYKRRKORAVMG 434

Db 794 K 794
Qy 435 K 435

RESULT 15
ID R15148 standard; Protein; 475 AA.
AC R15148;
DT 14-FEB-1992 (first entry)

Listing for Mary Hale

Tue Mar 19 12:36:56 1996

Page

Listing for Mary Hale

Tue Mar 19 12:36:56 1996

Page

DE Ro/SSA autoantigen.
 KW Autoantibody; autoantigen; SLE; systemic lupus erythematosus.
 OS Homo sapiens.
 PN W0911711-A.
 PD 14-NOV-1991.
 PF 07-MAY-1991; U03139.
 PR 07-MAY-1990; US-520270.
 PA (OKLA-) OKLAHOMA MED RES FO.
 PI Frank MB, Itoh K.
 DR WII; 91-3531248.
 DR N-FSDB; Q14798.
 PT DNA encoding an Ro-SSA autoantigen - useful for diagnosing
 auto-immune disorders or presence of auto-antibodies
 PS Disclosure; Fig 2; 41pp; English.
 CC A cDNA library (from human thymus mRNA) in lambda gt11 was screened
 with serum from a patient having systemic lupus erythematosus. Two
 clones were reactive with sera (from a panel of lupus patients)
 which contd. autoantibodies against 52 kd protein.
 CC both the cDNA and the protein expressed from it, or portions of it,
 are useful as diagnostic agents in the identification of patients
 having autoantibodies and in the identification and analysis of
 the structural and functional properties of the autoantigen and for
 application in immunotherapeutic regimens.
 CC Sequence 475 AA;

Query Match 2.6%; Score 107; DB 3; Length 475;
 Best Local Similarity 27.3%; Pred. No. 4; 70e+00;
 Matches 12; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

Db	16	cpicldpfpevpvsiqecghscqecisqvgkqggsvcpwqrqrfl	59
Qy	53	CEKCHLVLCSPROTECGHRFCESMAALLSSSPKCTACQESIV	96

Search completed: Tue Mar 19 09:53:45 1996
 Job time : 28 sec.

Post-processing: Minimum Match 0% Listing first 45 summaries

Database: pir46
 1:annl 2:ann2 3:ann3 4:unannl 5:unann2 6:unann3 7:unann4
 8:unann5 9:unann6 10:unann7 11:unann8 12:unrev2
 13:unrev2

Statistics: Mean 49.231; Variance 140.106; scale 0.351

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	4139	99.4	568	11	A55649 TNFR-associated prot	0.00e+00
2	812	19.5	416	11	B55649 TNFR-associated prot	2.29e-02
3	191	4.6	458	9	K29361 DG17 protein - slime	9.46e-09
4	164	3.9	760	11	A40195 meprin A (EC 3.4.24.	1.10e-05
5	149	3.6	283	5	A60364 tropomosin - migrat	2.89e-04
6	151	3.6	377	11	A67380 RING finger containi	2.89e-04
7	151	3.6	748	11	A24134 endopeptidase 2 (EC	2.89e-04
8	151	3.6	1669	2	S02771 myosin heavy chain A	1.23e-03
9	145	3.5	284	5	A4980 tropomyosin, oblique	1.23e-03
10	145	3.5	1959	2	A33977 myosin heavy chain,	1.23e-03
11	141	3.4	1961	2	A61231 myosin heavy chain N	1.23e-03
12	142	3.4	1966	2	M2NEW myosin heavy chain B	2.53e-03
13	132	3.2	284	5	S24972 tropomyosin alpha, c	2.67e-02
14	132	3.2	527	10	S33068 myosin II heavy chai	6.70e-02
15	132	3.2	700	10	S49383 meprin A (EC 3.4.24.	2.67e-02
16	135	3.2	714	13	S39464 PAB2 peptide hydrola	1.33e-02
17	135	3.2	979	2	S03166 myosin heavy chain,	1.33e-02
18	135	3.2	3559	10	S37536 macrogolgin - human	1.33e-02
19	130	3.1	221	5	S054946 tropomyosin, nonmus	4.24e-02
20	130	3.1	245	13	S24403 tropomyosin alpha -	4.24e-02
21	128	3.1	245	5	C34787 tropomyosin 3 alpha,	6.70e-02
22	128	3.1	248	5	D39816 tropomyosin 5b, fibr	6.70e-02
23	128	3.1	248	5	C39816 tropomyosin 5a, fibr	6.70e-02
24	128	3.1	251	5	B34787 tropomyosin 2 alpha,	6.70e-02
25	131	3.1	257	5	A02991 tropomyosin heavy chain,	3.37e-02
26	131	3.1	284	13	S23470 tropomyosin - Africa	3.37e-02
27	131	3.1	465	5	A02986 myosin alpha heavy c	3.37e-02
28	129	3.1	944	9	S26710 spindle pole body pr	5.33e-02
29	127	3.1	951	5	A33626 fibrinogen alpha cha	5.33e-02
30	129	3.1	1053	10	A31642 dyneactin - chicken	8.41e-02
31	127	3.1	1078	5	A30220 myosin heavy chain,	8.41e-02
32	1132	9	S37932 hypothetical protein	6.70e-02		

```

RESULT      1
ENTRY      A55649  #type complete
TITLE      TNFR-associated protein LAP1 - human
ORGANISM   CD40-binding protein
DATE      23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change
ACCESSIONS
REFERENCE
#authors Mosialos, G.; Birkenbach, M.; Yalamanchili, R.; VanArsdale,
T.; Ware, C.; Kieff, E.
#journal Cell (1995) 80:389-399
#title The Epstein-Barr virus transforming protein LMP1 engages
signaling proteins for the tumor necrosis factor receptor
family.

#accession A55649
#molecule_type mRNA
##residues 1-568 ##label MOS
##cross-references GB:U19260
##note nucleotide sequence not given

REFERENCE
#authors Hu, H.M.; O'Rourke, K.; Bouski, M.S.; Dixit, V.M.
#journal J. Biol. Chem. (1994) 269:30063-30072
#title A novel RING finger protein interacts with the cytoplasmic
domain of CD40.

#accession A55135
#molecule_type mRNA
##residues 1-133,135-404, 'G', 406-568 ##label HUA
##cross-references GB:U15637
##note nucleotide sequence not given
KEYWORDS
FEATURE
SUMMARY
53-91.

#length 568 #molecular_weight 64490 #checksum 8660

Query Match 99.4%; Score 4139; DB 11; Length 568;
Best Local Similarity 99.6%; Pred. No. 0-00e-00; Mismatches 566; Conservative 0; Indels 1; Gaps 1;
Db 1 messkkmndspalqtnppkltktldrsagtpvfpmeqggwykekfvktvedykckchlv 60
Qy 1 MESSKKNDSPGALQTNPPKLTKTLDRSAGTPVFPMEQGGYKEKFVKTVEDYKCKCHLV 60

```

```

RESULT      2
ENTRY      B55649  #type complete
TITLE      TNFR-associated protein EB16 - human
ORGANISM   #formal name Homo sapiens #common name man
DATE      23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change
ACCESSIONS
REFERENCE
#authors Mosialos, G.; Birkenbach, M.; Yalamanchili, R.; VanArsdale,
T.; Ware, C.; Kieff, E.
#journal Cell (1995) 80:389-399
#title The Epstein-Barr virus transforming protein LMP1 engages
signaling proteins for the tumor necrosis factor receptor
family.

#accession B55649
#status preliminary
#molecule_type mRNA
##residues 1-416 ##label MOS
##cross-references GB:U19261
SUMMARY
#length 416 #molecular_weight 46163 #checksum 6815
Query Match 19.5%; Score 812; DB 11; Length 416;
Best Local Similarity 38.9%; Pred. No. 2.29e-93;
```



```

RESULT      5
ENTRY
TITLE      A60364 #type complete
ORGANISM   tropomyosin - migratory locust
DATE       03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change
ACCESIONS
REFERENCE
#authors
Krieger, J.; Raming, K.; Knipper, M.; Grau, M.; Mertens, S.; Breer, H.
#Journal Insect Biochem. (1990) 20:173-184
#title Cloning, sequencing and expression of locust tropomyosin.
#accession A60364
#status not compared with conceptual translation
#molecule_type mRNA
#residues 1-283
#superfamily tropomyoain
#molecular_weight 32439 #checksum 4917
#length 283 #label KRT
KEYWORDS
SUMMARY
Query Match      3.5%; Score 149; DB 5; Length 283;
Best local Similarity 19.0%; Pred. No. 4.63e-04;
Matches 20; Conservative 36; Mismatches 46; Indels 3; Gaps 3;
Db      13 lekdnaldrallceggardanl-ra-ekaaearalqkqktiendldqtgeslgqvma 70
Qy      297 FEIEERQKEMRANSKILHQRIDSPAKKEKILDEIRPRQNMKAEDSMKSSVE 355
Db      71 leekkekallqnaesevalnrrqilederseerelatatakiae 115
Qy      356 LQNRVTELESVDKSAGQVARNTGILLESQLSRHDQMLSLVDIRLAD 400

```

```

RESULT      6
ENTRY
TITLE      A47380 #type complete
ORGANISM   RING finger-containing DNA binding protein RING1 - human
DATE       21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change
ACCESIONS
REFERENCE
#authors
Lovering, R.; Hanson, I.M.; Borden, K.L.; Martin, S.J.; O'Reilly, N.J.; Evan, G.I.; Reaman, D.; Pappin, D.J.; Trowsdale, J.; Freeman, P.S.
#journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:2112-2116
#title Identification and preliminary characterization of a protein motif related to the zinc finger.
#cross-reference MUID:93211912
#accession A47380
#status preliminary
#molecule_type protein
#residues 1-377 #label LOV
#cross-references NCBIN:128010; NCBIPI:128011
#note sequence extracted from NCBP backbone
SUMMARY
#length 377 #molecular_weight 39145 #checksum 7895
Query Match      3.6%; Score 151; DB 11; Length 377;
Best local Similarity 24.7%; Pred. No. 2.84e-04;
Matches 24; Conservative 30; Mismatches 39; Indels 4; Gaps 4;

```

```

RESULT      7
ENTRY
TITLE      S24134 #type complete
ORGANISM   endopeptidase 2 (EC 3.4.24.-) - rat
DATE       01-Sep-1995 #sequence_revision 01-Sep-1995 #text_change
ACCESIONS
REFERENCE
#authors
Corbeil, D.; Gaudoux, F.; Wainwright, S.; Ingram, J.; Kenny, A.J.; Boileau, G.; Crine, P.
#Journal FEBS Lett. (1992) 309:203-208
#accession S24134
#status preliminary
#molecule_type mRNA
#residues 1-748 #label COR
KEYWORDS
FEATURE
#binding_site zinc (H+) #status predicted
SUMMARY
#length 748 #molecular_weight 85138 #checksum 2333
Query Match      3.6%; Score 151; DB 11; Length 748;
Best local Similarity 28.9%; Pred. No. 2.84e-04;
Matches 26; Conservative 23; Mismatches 37; Indels 4; Gaps 4;
Db      436 vwtirnlnsqvlentvgdrl-v-sprfyse-ggfgvtlypngrisnsgylglafhly 492
Qy      418 IWRIDTKRKQEAIVNGKTISLYSPFVYFGKMCARVYLNDGMKGTHLSLFFVIM 477
Db      493 sgandvilewpprendamtildgedparn 522
Qy      478 RGYDALLWPWPKQKVITIM-LMDQSSRRH 506

```

```

RESULT      8
ENTRY
TITLE      S02771 #type complete
CONTAINS myosin heavy chain A - Caenorhabditis elegans
ORGANISM   myosin ATPase (EC 3.6.1.32)
DATE       31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
27-Jan-1995
ACCESIONS
REFERENCE
#authors
#Journal Dibb, N.J.; Maruyama, I.N.; Krause, M.; Karn, J. J. Mol. Biol. (1999) 205:603-613
#title Sequence analysis of the complete Caenorhabditis elegans

```

myosin heavy chain gene family.

```

#cross-references MUID:89178677
#accession S0271
##molecule_type DNA
##residues 1-1969 ##label DIB
##note the nucleotide sequence is not given

GENETICS
#gene myo-3
#introns 46/1; 116/3; 169/1; 269/1; 445/2; 1898/3
#superfamily myosin heavy chain; myosin head homology
CLASSIFICATION
KEYWORDS
FEATURE
89-779
179-186
667-689
770-784
852-1969
852-1166
852-1166
1167-1969
130
#domain myosin head homology ##label HEA \
#region nucleotide-binding motif A (P-loop) \
#region actin-binding #status predicted \
#region actin-binding #status predicted \
#domain coiled coil ##status predicted ##label COT \
#region S2 \
#region light meromyosin \
#modified site N6,N6,N6-trimethyllysine (Lys) ##status
predicted \
#binding site ATP (Lys) ##status predicted \
#active site Cys ##status predicted ##checksum 6892
SUMMARY
#length 1969- ##molecular-weight 225508 ##checksum 6892

Query Match
Best Local Similarity 3.5%; Score 145; DB 5; Length 284;
Matches 26; Conservative 49; Mismatches 41; Indels 6; Gaps 6;
Query Match
Best Local Similarity 21.3%; Score 123e-03; DB 2; Length 284;
Matches 26; Conservative 49; Mismatches 41; Indels 6; Gaps 6;
Db 985 dunirsldqndehaqavakinket-kheasnkrknledlqseedkvnhlelkirknleqq 1043
Qy 270 EKKVSLQNESVEKNSQIOSHNLQCSFEIEERO-KEMLRNNESKILHLQRVIDSAER 328
Db 1044 mdeeleenidirekrksdgikakrkrqg-ekvqaq-enaeitcq--th-dv-ettikrke 1097
Qy 329 LKEIDKEIRPFKRNWEADSKMSSVSLQNRVTELESVDKSAGQVARNTGILSQLSRHD 388
Db 1098 edalhntnaklae 1109
Qy 389 QMLSVHDIRLAD 400

RESULT 9
ENTRY 10
TITLE A33977 ##type complete
CONTAINS myosin heavy chain, cellular - chicken
ORGANISM myosin ATPase (EC 3.6.1.32)
DATE 31-Dec-1993 ##sequence_revision 31-Dec-1993 ##text_change
27-Jan-1995
ACCESSIONS A33977
REFERENCE A33977
#authors Shonet, R.V.; Conti, M.A.; Kawamoto, S.; Preston, Y.A.; Brill, D.A.; Adelstein, R.S.
Proc. Natl. Acad. Sci. U.S.A. (1989) 86:7726-7730
#journal Cloning of the cDNA encoding the myosin heavy chain of a
#title vertebrate cellular myosin.
#cross-references MUID:90046668
#accession A33977
#molecule_type mRNA
##residues 1-1959 ##label SHO
##cross-references GB:M28510
CLASSIFICATION
KEYWORDS
#superfamily myosin heavy chain; myosin head homology
actin binding; ATP; coiled coil; hydrolase; methylation;
tandem repeat

FEATURE
89-764
174-181
552-565
626-640
837-1936
837-1277
1278-1959
1937-1959
125-
#domain myosin head homology ##label HEA \
#region nucleotide-binding motif A (P-loop) \
#region actin-binding #status predicted \
#region actin-binding #status predicted \
#domain coiled coil ##status predicted ##label COT \
#region S2 \
#region light meromyosin \
#domain carboxyl-terminal ##label CBT \
#modified site N6,N6,N6-trimethyllysine (Lys) ##status
predicted \
#binding site ATP (Lys) ##status predicted \
#active site Cys ##status predicted ##checksum 3641
#accession A44980
#status preliminary
#molecule_type mRNA
SUMMARY
#length 1959- ##molecular-weight 226502 ##checksum 3641
Query Match
Score 145; DB 2; Length 1959;
```

Best local Similarity 25.2%; Pred. No. 1.23e-03; Matches 34; Conservative 34; Mismatches 57; Indels 10; Gaps 8; Db 1460 yaerdraeaearaketkalslaraleeinqkaelerlnqfr-t--emelmskadv 1516
 Qy 297 FEIEIERQEMIRNNEKILHQLRVIDSQAEKIKEDEIRPRQNEEAESMKSSVEL 356
 Db 1517 gksvhelakrakeqveemtkqe-e--edelqatedaklri-levnqgam-kqfr 1571
 Qy 357 QNRVTELESVDKS-ACQVARNTGLLESQI-SRHDOMLSVHDIRLADMIGFQVLETASYN- 414
 Db 1572 dligrdeeqneekrkq 1586
 Qy 415 GYLIWKIRDYKRKO 429

RESULT 11
 ENTRY A61231 #type complete
 TITLE myosin heavy chain NM_002101
 ORGANISM myosin ATPase (EC 3.6.1.32)
 DATE 27-Jan-1995 #sequence_revision 14-Jul-1994 #text_change
 ACCESSIONS A61231; A34876
 REFERENCE #authors Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yamakawa, K.; Guly, D.; Adelstein, R.S.; Weir, L.
 #journal Circ. Res. (1991) 69:530-539
 #title Human nonmuscle myosin heavy chains are encoded by two genes
 #accession A61231
 #molecule_type mRNA
 #residues 1-715 #label SIM
 #cross-references GB:W69180
 REFERENCE A34976
 #authors Saez, C.G.; Myers, J.C.; Shows, T.B.; Leinwand, L.A.
 Proc. Natl. Acad. Sci. U.S.A. (1990) 87:1164-1168
 #title Human nonmuscle myosin heavy chain mRNA: generation of diversity through alternative polyadenylation.
 #cross-references MUID:90138958
 #accession A34976
 #molecule_type mRNA
 #residues 715-1961 #label SAE
 #cross-references GB:M31013
 GENETICS
 #gene GDB:MYH9
 #map Position 22q12.3-q13.1
 CLASSIFICATION #superfamily myosin heavy chain; myosin head homology; actin binding; ATP; coiled coil; hydrolase; methylation; tandem repeat
 FEATURE 84-764
 #domain myosin head homology #label HEA\\
 #region nucleotide-binding motif A (P-loop)\\
 #region actin-binding #status predicted\\
 #region actin-binding #status predicted\\
 #domain coiled coil #status predicted #label COT\\
 #region S2\\
 #region light meromyosin\\
 #domain carboxyl-terminal #label CBT\\
 #modified_site N6,N6,N6-trimethyllysine (Lys) #status
 125 1166-1966

180 #binding site ATP (Lys) #status predicted\\
 694, 704 #actin_gite Cys #status predicted
 SUMMARY #length 1961-#molecular-weight 226741 #checksum 9108
 Query Match 3.4%; Score 141; DB 2; Length 1961;
 Best local Similarity 27.8%; Pred. No. 3.21e-03; Matches 32; Conservative 26; Mismatches 48; Indels 9; Gaps 7;
 Db 1461 yaerdraeaearaketkalslaraleeameqkaelerlnqfr-t--emelmskadv 1517
 Qy 297 FEIEIERQEMIRNNEKILHQLRVIDSQAEKIKEDEIRPRQNEEAESMKSSVEL 356
 Db 1518 gksvhelakrakeqveemtkqe-e--edelqatedaklri-levnqgam-kqfr 1571
 Qy 357 QNRVTELESVDKS-ACQVA-RNTGL-LESQI-SRHDOMLSVHDIRLADMIGFQ 406

RESULT 12
 ENTRY MWK9 #type complete
 TITLE myosin heavy chain B - Caenorhabditis elegans
 ORGANISM myosin ATPase (EC 3.6.1.32)
 DATE 13-Jun-1983 #sequence_revision 19-Feb-1984 #text_change
 ACCESSIONS A93958; A93287; A02992
 REFERENCE #authors Karni, J.; Brenner, S.; Barnett, I.
 Proc. Natl. Acad. Sci. U.S.A. (1983) 80:4253-4257
 #title Protein structural domains in the *Caenorhabditis elegans* unc-54 myosin heavy chain gene are not separated by introns.
 #cross-references MUID:83273600
 #accession A93958
 #molecule_type DNA
 #residues 1-966 #label KAR
 #cross-references GB:J01050
 REFERENCE A93287
 #authors McLachlan, A.D.; Karn, J.
 Nature (1982) 299:226-231
 #journal Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle.
 #cross-references MUID:82272395
 #accession A93287
 #molecule_type RNA
 #residues 850-1336, 'R', 1338-1879, 'L', 1881-1966 ##label MCL
 GENETICS
 #gene unc-54
 #map Position 21/3; 64/2; 114/3; 267/1; 528/3; 1750/3; 1822/3; 1897/3
 CLASSIFICATION #superfamily myosin heavy chain; myosin head homology
 KEYWORDS actin binding; ATP; coiled coil; hydrolase; methylation; muscle contraction; tandem repeat
 FEATURE 84-764
 #region nucleotide-binding motif A (P-loop)\\
 #region actin-binding #status predicted\\
 #region actin-binding #status predicted\\
 #region coiled coil #status predicted #label COT\\
 #region S2\\
 #region light meromyosin\\
 #domain carboxyl-terminal #label CBT\\
 #modified_site N6,N6,N6-trimethyllysine (Lys) #status
 125 1166-1966
 #region light meromyosin\\

183 #modified site N6,N6,N6-trimethyllysine (Lys) #status predicted\\
 705,715 #binding-site ATP (Lys) #status predicted\\
 SUMMARY #active-site Cys #status predicted
 Query Match #length 1966 #molecular-weight 225125 #checksum 6382
 Best Local Similarity 24.2%; Pred. No. 2.53e-03; Mismatches 37; Indels 8; Gaps 8;
 Matches 32; Conservative 37; MisMatches 55; Indels 8; Gaps 8;
 Db 984 dbqrlslqdeaklkkkek-khqeelnrkimqdeekgqhnqkvakleqt 1042
 Qy 270 EKKVSLQNEVEKURSKIQSFLNQICSEIEEROKEMLRNNESKLHLQRVIDSQAEKL 328
 Qy 329 IKLDEKEIRPRFQNWEEADSKSSVSLQNRTVTELESVDKSAGQVARNTGILLESQSLRD 388
 Db 1097 selfhassled 1108
 Qy 389 QMLSVHDIDRLAD 400

RESULT 13
 ENTRY S24972 #type complete
 TITLE tropomyosin alpha, cardiac - pig
 ORGANISM #formal name Sus scrofa domestica #common name domestic pig
 DATE 12-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
 ACCESSIONS S24972
 REFERENCE S24972
 #authors Whitby, F.G.; Kent, H.M.; Stewart, F.; Stewart, M.; Xie, X.; Hatch, V.; Cohen, C.; Phillips Jr., G.
 #submission submitted to the EMBL Data Library, April 1992
 #description Structure of tropomyosin at 9 Angstroms resolution.
 #classification #superfamily S24972
 #accession #status preliminary
 #molecule_type mRNA
 #residues 1-284 #label WHI
 #cross-references EMBL:X66274
 #description #superfamily tropomyosin
 SUMMARY #length 284 #molecular-weight 32729 #checksum 9163
 Query Match #length 3 3.2%; Score 132; DB 5; Length 284;
 Best Local Similarity 23.9%; Pred. No. 2.67e-02; Mismatches 27; Indels 8; Gaps 6;
 Matches 27; Conservative 39; MisMatches 39; Indels 8; Gaps 6;
 Db 45 slqkkkkatedelkysalkdakqeklaekkataed-vasnrrqlf-e-eeld 100
 Qy 289 SLRQNCSPFELTEROKEMLRNNESKL-IHLQRVIDSOAERKELKEIRPRFQNWEEAD 347
 Db 101 --raq-erlatlalqklegaaekadesergmkviesrckqdkmekmeiqekle 150
 Qy 348 SMKSSVESQNRTELESVDSKAGQVARNTGILLESQSLRDHOMLSVHDIDRLAD 400

RESULT 14
 ENTRY S33068 #type complete
 TITLE myosin II heavy chain - fluke (Schistosoma mansoni)
 ORGANISM #formal_name Schistosoma mansoni

ACCESSIONS S33068
 REFERENCE A46514
 #authors Soisson, L.M.A.; Masterson, C.P.; Tom, T.D.; McNally, M.T.; Lowell, G.H.; Strand, M.
 DATE 26-May-1995
 SUMMARY #journal J. Immunol. (1992) 149:361-3620
 #title Induction of protective immunity in mice using a 62-kDa recombinant fragment of a Schistosoma mansoni surface antigen.
 #cross-references MUID:93056536
 #accession S33068
 #status preliminary
 #molecule_type mRNA
 #residues 1-527 #label SOI
 #cross-references EMBL:X65591
 #note the authors translated the codon CAA for residue 346 as Ly8

RESULT 15
 ENTRY S49383 #type complete
 TITLE mephrin A (EC 3.4.24.18) - human
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 16-May-1995 #sequence_revision 12-May-1995 #text_change
 ACCESSIONS S49383
 REFERENCE S49383
 #authors Elderling, J.A.; Groenberg, J.; Sterchi, E.E.
 #submission submitted to the EMBL Data Library, September 1994
 #description Cloning and the PABA-Peptide hydrolase beta subunit: coexpression is required for plasma membrane localization of the alpha subunit in COS-1 cells.
 #accession S49383
 #status preliminary
 #molecule_type mRNA
 #residues 1-700 #label ELD
 #cross-references EMBL:X81333
 CLASSIFICATION #superfamily MAM homology; astacin homology
 FEATURE 71-257
 #domain astacin homology #label AST
 #length 700 #molecular-weight #label MAM
 SUMMARY #length 700 #molecular-weight #label MAM
 #checksum 916

Query Match #length 3 3.2%; Score 132; DB 10; Length 700;
 Best Local Similarity 33.3%; Pred. No. 2.67e-02; Mismatches 30; Indels 11; Gaps 7;
 Matches 20; Conservative 29; MisMatches 29; Indels 11; Gaps 7;

Listing for May Hale
10a Mar 1912-36-6/1998

22

LITERACY IN THE PRIMARY GRADES / 1550

(TM)                   

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.

THE JOURNAL OF CLIMATE

Run On: Tue Mar 19 09:51:11 1996; MsPar time 10.24 Seconds
MSDBL_1p protein - protein database search, using BLASTN algorithm

Tabular output not generated.

Title: >US-08-404-832-2

perfect score: 4162

Scoring table: PAM 150

卷之三

SEARCHED. INDEXED. SERIALIZED. FILED.

post-processing. Millions of batches of first 45 summaries

Database: swissprot31

1 part 2 parts 3 parts 4 parts 5 parts 6 parts 7 parts

Statistics: Mean 51.439; Variance 108.497; scale 0.474

ALIGNMENTS

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ID TRF2 MOUSE STANDARD; PRT; 501 AA.
AC P39429;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE TNF RECEPTOR ASSOCIATED FACTOR 2 (TRAF2).
OS MUSCULUS (MOUSE).

ALIGNMENTS

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1] SEQUENCE FROM N.A.
 RP RM 94349371
 RA ROTHE, M., WONG, S.C., HENZEL, W.J., GOEDDEL, D.V.;
 RL CELL, 78: 681-692 (1994).
 CC -1- FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN
 OF THE 75 KD TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
 DR EMBL; L35303; MMTR2A.
 KW ZINC-FINGER; COILED COIL.
 FT ZN FING 34 72 C3HC4-TYPE.
 SQ SEQUENCE 501 AA; 56026 MW; 1264825 CN;
 Query Match 19.2%; score 798; DB 7; Length 501;
 Best Local Similarity 52.7%; Pred. No. 5.03e-123;
 Matches 98; Conservative 49; Mismatches 37; Indels 2; Gaps 2;
 Db 317 lsnkvqglertigkldamadleckvselvestvdtyagvfiwksdttkrqavagrtpai 376
 Qy 380 LESQLSRHDMLSVHDRLADMDLGFOVLETASNGVILWIKDVKRQBAVMKTLS 439
 Db 377 fspafyftrsgykmclrlvngtqgrghisffwvngkqpdndlqwpfqkvmlid 436
 Qy 440 YSQPFITGYGKMCARVYANGDGKGKGHLSLPVIDGEVDALIWPFKOKVITIM 499
 Db 437 h-nrrevhvdafrpvtsssfqrpsdmiasgplfpvskmeaksyvrdaiikai 495
 Qy 500 QGSSRRHLGDKAPDPNSSFKKPTGEMNIASGCPVFAVQVLE-NGTYIKDTIFIKV 558
 Db 496 vdltg1 501
 Qy 559 VDTSDL 564

RESULT 2
 ID TNF1 MOUSE STANDARD; PRT; 409 AA.
 AC P39428;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE TNF RECEPTOR ASSOCIATED FACTOR 1 (TRAFL).
 OS MOS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 123-135 AND 390-402.
 RM 94349371
 RA ROTHE, M., WONG, S.C., HENZEL, W.J., GOEDDEL, D.V.;
 RL CELL, 78: 681-692 (1994).
 CC -1- FUNCTION: SIGNAL TRANSDUCER ASSOCIATED WITH THE CYTOPLASMIC DOMAIN
 OF THE 75 KD TUMOR NECROSIS FACTOR RECEPTOR (TNF-R2).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SUBUNIT: HETERODIMER OF TRAF1 AND TRAF2.
 DR EMBL; L35302; MMTR2A.
 KW COILED COIL.
 SQ SEQUENCE 409 AA; 45464 MW; 857796 CN;

RESULT 3
 ID DG17 DICDI STANDARD; PRT; 458 AA.
 AC P11467;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE DG17 PROTEIN.
 GN ZFAA OR DG17.
 OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
 OC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA; RHIZOPODA;
 OC DICTYOSTELOMATA.
 RN [1] SEQUENCE FROM N.A.
 RP DRISCOLL, D.M., WILLIAMS, J.G.;
 RM DRISCOLL, D.M., WILLIAMS, J.G.;
 RL MOL. CELL. BIOL. 7:4482-4489 (1987).
 CC -1- THE EXPRESSION OF DG17 PROTEIN IS DEVELOPMENTALLY REGULATED.
 CC -1- INDUCTION: BY CAMP DURING AGGREGATION.
 DR EMBL; M18106; DDG17A.
 PIR; A29361; A29361.
 DR DICTYDB; DD02010; ZFAA.
 KW DEVELOPMENTAL PROTEIN; ZINC-FINGER.
 FT SIMILAR 25 67 TO DROSOPHILA SINA (AA 71-108).
 FT ZN FING 27 66 POTENTIAL.
 FT ZN FING 178 198 POTENTIAL.
 SQ SEQUENCE 458 AA; 53015 MW; 1107348 CN;

CC	Query Match	4.6%	Score 191;	DB 2;	Length 458;
CC	Best Local Similarity	25.8%	Pred. No. 1, 98e-12;		
CC	Matches	39;	Conservative	36;	Mismatches 65;
CC			Indels	11;	Gaps 10;
CC					
Db	124	kdeengckekievqgqidlin-cqykvfcfkfq--ekirr--m-nskmgqfkivc	178		
Qy		:: : : ; :: .: : : : ; : : :			
Db	118	RNESRGCAEQLQMLGH-VHLKNDCHFEEFLPCVPRDPCKEVRLRDVERACKREATC	176		
Qy		:: : : ; :: .: : : : ; : : :			
Db	179	dfckddkkkelehykt-cpmpvdcsgqsvk-iervkidihiendctqpcyf	236		
Qy		:: : : ; :: .: : : : ; : : :			
Db	177	SHCK-SQVNIALQKHDPEPDCPVCVNSPHKSQVTSRISHSAHL-SECNARSTCSRK	234		
Qy		:: : : ; :: .: : : : ; : : :			
Db	237	eqgckvemkselqhlervnhnpynglie	267		
Qy		:: : : ; :: .: : : : ; : : :			
Db	235	RYGGVFQGTMQQIKAHEASSAVQH-VNLKE	264		
Qy		:: : : ; :: .: : : : ; : : :			
RESULT	4				
ID	MEPA_MOUSE	STANDARD;	PRT;	760 AA.	
AC	P28825;				
DT	01-DEC-1992	(REL. 24, CREATED)			
DT	01-DEC-1992	(REL. 24, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)			
DE	MEPRIN A ALPHA-SUBUNIT PRECURSOR (EC 3.4.24.18) (ENDOPEPTIDASE-2) (MEP-1).				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; EUTHERIA; RODENTIA.				
RN	[1]	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RP	STRAIN=C57BL/6, AND C3H/HE; TISSUE=KIDNEY;				
RC	92250517				
RA	JIANG W., GORBEA C.M., FLANNERY A.V., BEYNON R.J., GRANT G.A., BOND J.S.;				
RA	J. BIOL. CHEM. 267:9185-9193(1992).				
RN	[2]	SEQUENCE OF 77-275 FROM N.A.			
RP	92040208				
RA	DUMERUTH E., STERCHI E.J., JIANG W., WOLZ R.L., BOND J.S., FLANNERY A.V., BEYNON R.J., GORBEA C.M., BOND J.S.;				
RA	J. BIOL. CHEM. 266:21381-21385(1991).				
RN	[3]	CHARACTERIZATION.			
RP	9135206				
RA	WOLZ R.L., HARRIS R.B., BOND J.S.;				
RA	BIOCHEMISTRY 30:8488-8493(1991).				
RL	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEIN AND PEPTIDE SUBSTRATES PREFERENTIALLY ON CARBOXYL SIDE OF HYDROPHOBIC RESIDUES.				
CC	-1- COFACTOR: ZINC.				
CC	-1- SUBUNIT: HOMOTETRAMER OF ALPHA OR BETA SUBUNITS; HETEROCTETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS ARE FORMED BY NON-COVALENT ASSOCIATION OF TWO DISULFIDE-LINKED HETERODIMERS; GENETIC FACTORS DETERMINE WHICH OLIGOMER(S) WILL BE FORMED (STRAIN-SPECIFIC).				
CC	-1- PTM: N-GLYCOSYLATED; AT LEAST 3 OF THE POTENTIAL SITES ARE USED.				
CC	-1- TISSUE SPECIFICITY: KIDNEY, INTESTINAL BRUSH BORDERS, AND SALIVARY DUCTS.				
CC	-1- SIMILARITY: HIGH TO OTHER ZINC METALLOPROTEASES. BELONGS TO THE ASTACIN SUBFAMILY.				
RESULT	5				
ID	TPMM_L0CM1	STANDARD;	PRT;	283 AA.	
AC	P31816;				
DT	01-JUL-1993	(REL. 26, CREATED)			
DT	01-JUL-1993	(REL. 26, LAST SEQUENCE UPDATE)			
DT	01-JUL-1993	(REL. 26, LAST ANNOTATION UPDATE)			
DE	TROPOMYOSIN, MUSCLE.				
OS	LOCUSTA MIGRATORIA (MIGRATORY LOCUST).				
CC	--!- SIMILARITY: THE PROTEIN INCLUDES 1 EGF-LIKE REPEAT.				
CC	-!- SIMILARITY: CONTAINS A MAM DOMAIN.				
DR	DR PROSITE; PS00142; ZINC_PROTEASE.				
DR	DR PROSITE; PS00740; MAM.				
DR	KW HYDROLASE; METALLOPROTEASE; ZINC; GLYCOPROTEIN; TRANSMEMBRANE.				
DR	KW ZYMOGEN; SIGNAL; EGF-LIKE DOMAIN; 3D-STRUCTURE.				
FT	SIGNAL	1		33	
FT	PROPEL	34		77	
FT	CHAIN	78		760	
FT	DOMAIN	78		726	
FT	TRANSEM	727		754	
FT	DOMAIN	755		760	
FT	DOMAIN	78		275	
FT	REPEAT	276		445	
FT	METAL	685		725	
FT	ACT SITE	167		167	
FT	METAL	168		168	
FT	METAL	171		171	
FT	DISULFID	177		177	
FT	DISULFID	688		699	
FT	DISULFID	693		708	
FT	DISULFID	710		723	
FT	CARBOHYD	41		41	
FT	CARBOHYD	152		152	
FT	CARBOHYD	234		234	
FT	CARBOHYD	270		270	
FT	CARBOHYD	330		330	
FT	CARBOHYD	426		426	
FT	CARBOHYD	452		452	
FT	CARBOHYD	545		546	
FT	CARBOHYD	553		553	
FT	CARBOHYD	614		614	
FT	CARBOHYD	614		614	
FT	CONFLICT	272		272	
FT	SQ	760 AA;		85702 MW;	T -> A (IN REF. 2).
Qy	SEQUENCE	760 AA;		2920193 CN;	
Query Match	3.9%	Score 164;	DB 5;	Length 760;	
Best Local Similarity	30.0%	Pred. No. 1, 98e-08;			
Matches	271;	Conservative	25;	Mismatches 34;	Indels 4;
					Gaps 4;
Db	447	wvtirnqisilentvkagkl-v-sprfyse-qygvgvtlypngritsnsglglftly	503		
Qy		:: : ; :: : : ; : : ; : : ; : : ; : : ; : :			
Db	418	IWKTRDKRKRRKQKVKMGTKLTSVSPPFYGTGKQCARVYLNGDGMGRKTHISLFFVIM	477		
Qy		:: : ; :: : : ; : : ; : : ; : : ; : : ; : :			
Db	504	sgndaiialewpvnraqaintldqeadrn	533		
Qy		:: : : : : ; :: : : ; : : ; : : ; : : ; : :			
Qy	478	RGEVDALLWPFKVKTLMQDGSSRRH	506		

OC EURYKOTA; METAZOA; ARTHROPODA; INSECTA; ORTHOPTERA.
 RN [1] SEQUENCE FROM N.A.
 RP RA KRIEGER J., RAMING K., KNIPPER M., GRAU M., MERTENS S., BREER H.;
 RL INSECT BIOCHEM. 20:173-184 (1990).
 CC -!- FUNCTION: TROPOMYOSIN, IN ASSOCIATION WITH THE TROPONIN COMPLEX,
 PLAYS A CENTRAL ROLE IN THE CALCIUM DEPENDENT REGULATION OF
 CC MUSCLE CONTRACTION.
 CC -!- DOMAIN: THE MOLECULE IS IN A COILED COIL STRUCTURE. THE SEQUENCE
 CC EXHIBITS A PROMINENT SEVEN-RESIDUES PERIODICITY.
 DR PIR; A60364; A60364.
 DR PROSITE; PS00325; Cofilin TROPOMYOSIN.
 DR PROSITE; PS00326; TROPOMYOSIN.
 KW MUSCLE PROTEIN; COILED COIL; REPEAT;
 SQ SEQUENCE 283 AA; 32439 MW; 334140 CN;

Query Match 3.6%; Score 149; DB 7; Length 283;
 Best Local Similarity 19.0%; Pred. No. 2.69e-06;
 Matches 20; Conservative 36; Mismatches 46; Indels 3; Gaps 3;

Db 13 leekakallqhaesevaanrriqleeldeleerseerlatataklae 115
 :|: ;|: ::|: ;|: |::|: ;|: ;|: ;|: ;|: ;|: ;|:
 Qy 297 FEIETEROKEMLRNNEKSHIKLORVIDSPAEKLEIDKEIRPFRNWEA-DSMKSSVES 355

Db 71 leekakallqhaesevaanrriqleeldeleerseerlatataklae 115
 :|: ;|: ::|: ;|: |::|: ;|: ;|: ;|: ;|: ;|:
 Qy 356 LQNRVTELESVDKSAGQVWNTGLESQSRHQMISVHDIRAD 400

RESULT 6
 ID RNN1 HUMAN STANDARD; PRT; 377 AA.
 AC 006587;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE RING1 PROTEIN.
 GN RING1.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1] SEQUENCE FROM N.A.
 RP RM 932211912
 RA LOVERING R., HANSON I.M., BORDEN K.L.B., MARTIN S., O'REILLY N.J.,
 RA EVAN G.I., RAHMAN D., PAPPIN D.J.C., TROWBRIDGE J., FREEMONT P.S.;
 RA RNL NATL. ACAD. SCI. U.S.A. 90:2112-2116(1993).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -!- SIMILARITY: CONTAINS A C3HC4-C-CLASS ZINC FINGER.
 DR EMBL; Z14000; HSRING1.
 DR PIR; A47380; A47380.
 DR HSSP; P28900; ICHC.
 DR PROSITE; PS00516; ZINC FINGER C3HC4.
 KW ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.
 FT ZN FING 19 58 C3HC4-TYPE.
 FT DOMAIN 176 231 GLY-RICH.
 FT DOMAIN 285 348 GLY-RICH.
 FT DOMAIN 172 175 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 377 AA; 39145 MW; 716288 CN;

Query Match 3.6%; Score 151; DB 6; Length 377;

Best Local Similarity 24.7%; Pred. No. 1.41e-06;
 Matches 24; Conservative 30; Mismatches 39; Indels 4; Gaps 4;

Db 11 relselmcpiclndmkntktckehrfosdcivtalsrgnhecptorkklvskslrp 70
 ::|: ;|: |::|: ;|: |::|: ;|: |::|: ;|: |::|: ;|: |::|:
 Qy 45 KVVEDKYKCEKHICLICSQRT-ECHRCESCEMAALISSSPKCTACQESIVKDKFQ 103

Db 71 dp-nfballisklyppreyeeahgqvrlr-lrlsrhng 105
 :|: ;|: |::|: ;|: |::|: ;|: |::|: ;|: |::|:
 Qy 104 NCCKRELLAQY-CRNESRGAEGQIMLGHVLHKND 139

RESULT 7
 ID MUSA CEEBL STANDARD; PRT; 1969 AA.
 AC PI2874;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
 DE MYSIN HEAVY CHAIN A (MHC A).
 GN MYO-3.
 OS GLONDRABBITIS ELEGANS.
 OC EUKARYOTA; METAZOA; ACELEMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
 RN [1]
 SQ SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RM 89178677
 RA DIBB N.J., MARTYAMA J.N., KRAUSE M., KARN J.;
 RL J. MOL. BIOL. 205:603-613 (1989).
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM)
 CC AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATTER BE SPLIT FURTHER INTO
 CC 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).
 CC -!- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -!- SIMILARITY: THE GLOBULAR HEAD AA SEQUENCE SHOWS A HIGH DEGREE OF
 CC SIMILARITY WITH THE GLOBULAR HEAD SEQUENCES OF MUSCLE & NONMUSCLE
 CC HEAVY CHAINS. BY CONTRAST, THE ROD SEQUENCE IS LESS CONSERVED, BUT
 CC THE PERIODICITIES OF HYDROPHOBIC & CHARGED RESIDUES, WHICH DICTATE
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -!- THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C. ELEGANS.
 CC -!- MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY WALL MUSCLE.
 CC THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 DR EMBL; X08067; CEMYO3.
 DR PIR; S02771; S02771.
 DR HSSP; P24733; ICM.
 KW MYOSIN; MUSCLE PROTEIN; COILED COIL; THICK FILAMENT; ACTIN-BINDING;
 KW ATP-BINDING; METHYLATION; ALKYLATION; HEPTAD REPEAT PATTERN;
 KW MULTIGENE FAMILY.
 FT DOMAIN 1 851 GLOBULAR HEAD (S1).
 FT DOMAIN 1 852 1969 ROD-LIKE TAIL (S2 AND LMM DOMAINS).
 FT NP-BIND 179 186 ATP.
 FT DOMAIN 667 689 ACTIN-BINDING.
 FT DOMAIN 770 784

Listing for Mary Hale

१

Listing for Mary Hale

Tue Mar 19 12:36:57 1996

	Query	Match	Score	Length	Pred.	No.	Best Local Similarity	21.2%	FT	MOD_RES	130	130	METHYLATION ('TRI-') (POTENTIAL).			
							Matches	28;	FT	MOD_RES	707	707	ALKYLATION (SH-1) (POTENTIAL).			
							Conservative	49;	FT	MOD_RES	717	717	ALKYLATION (SH-2) (POTENTIAL).			
	Qy	Qy	Qy	Qy	Qy	Qy	Qy	Qy	FT	SEQUENCE	1969	AA;	225509	MW;	15524438	CN;
RESULT	8								FT							
ID	MYSN_CHICK								FT	STANDARD;						
AC	P14105;								FT	PRT;	1959	AA.				
DT	01-JAN-1990	(REL.	13,	CREATED)					FT							
DT	01-JAN-1990	(REL.	13,	LAST	SEQUENCE	UPDATE)			FT							
DT	01-JUN-1994	(REL.	29,	LAST	ANNOTATION	UPDATE)			FT							
DE	MYOSIN HEAVY CHAIN, NONMUSCLE (NMHC).								FT							
OS	GALLUS GALLUS (CHICKEN).								FT							
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE; GALLIFORMES.								FT							
RA	SHOHEY R.V., CONTE M.A., KAWAMOTO S., PRESTON Y.A., BRILL D.A., ADDISSTEIN R.S., PROC. NATL. ACAD. SCI. U.S.A. 86:7726-7730(1989).								FT							
RA	PROC. NATL. ACAD. SCI. U.S.A. 86:7726-7730(1989).								FT							
RP	SEQUENCE FROM N.A. TISSUE-INTESTINAL EPITHELIUM;								FT							
RC	90046668								FT							
CC	SHOHEY R.V., CONTE M.A., KAWAMOTO S., PRESTON Y.A., BRILL D.A., ADDISSTEIN R.S., PROC. NATL. ACAD. SCI. U.S.A. 86:7726-7730(1989).								FT							
CC	REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).								FT							
CC	-!- EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MERMYSIN (LM) AND 1 HEAVY MERMYSIN (HM). IT CAN LATTER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).								FT							
CC	-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHAI HELICAL COILED COILS.								FT							
CC	-!- SIMILARITY: BELONGS TO THE MYOSIN HEAVY CHAIN FAMILY. STRONGEST TO OTHER NONMUSCLE MYOSINS.								FT							
DR	EMBL; M2651; GEMTHN.								FT							
DR	PIR; A33977; A33977.								FT							
DR	HSSP; P24733; 1SCM.								FT							
KW	MYOSIN; COILED COIL; ACTIN-BINDING; ALKYLATION; ATP-BINDING; HETATD REPEAT PATTERN; MULITIGENE FAMILY.								FT							
DOMAIN	1	835							FT	GLOBAL HEAD (S1);						
DOMAIN	836	1926							FT	RODLIKE TAIL (S2 AND LM DOMAINS).						

	Query Match	Score 3.5%;	DB 174; length 1959;
	Best Local Similarity	Matched 25.2%;	Length 1959;
	Local Similarity	Score 9.39%;	Length 1959;
FT	NP BIND	181	ATP.
FT	DOMAIN	654	ACTIN-BINDING.
FT	MOD RES	694	ALKYLATION (SH-1) (POTENTIAL).
FT	MOD RES	704	ALKYLATION (SH-2) (POTENTIAL).
SQ	SEQUENCE	1959 AA;	15394260 CN;
Db	1460 yaedeareearektaalslarlaeakqkalerunqfr-t--emedlmsekdv 1516	Qy	297 FEIEIERKEMURNNEKSILHQVQIDSQEKELDEIRPFRQNTEADMSSKSSVSL 356
Db	1517 gksvhelakralegqeveenkqie-el--edlqatedakir-levnqga-kaqfdr 1571	Qy	357 QNRVTELESVDKS-AGQVARNITGLESOLSRHDOMLSVHDRIADMDLGFOVLETASYN- 414
Db	1572 dligrdegneeqrkq 1586	Qy	415 GVLWKIRDKYKRKQ 429

Listing for Mary Hale

QY ||| : | | : | : :: : | : : | : | | : | : |
 318 LQRVIDSQAEKIKEL-DKEIRPFP-RQNWEADSKSSVLSQNRVTELESVDKSAGQVAR 375
 Db 90 gmkviesraqkdeekmeiqike 114
 Qy ::||: ; :::: ::||: ;
 376 NTGLLESQLSRHDQMLSVHDIRLAD 400

RESULT 13
 ID TPMY RAT STANDARD; PRT; 251 AA.
 AC P18343;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
 DE TROPOMYOSIN ALPHA CHAIN, BRAIN-2 (TMBR-2).
 OS ALPHATM.
 OC RAT(TUS NORVEGICUS (RAT)).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC ETHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N A.
 RC TISSUE=BRAIN;
 RM 90205854
 RA LEES-MILLER J.P.; GOODWIN L.O.; HELFMAN D.M.;
 RL MOL. CELL. BIOL. 10:1729-1742(1990).
 CC -!- FUNCTION: THE FUNCTION OF TROPOMYOSIN IN SMOOTH MUSCLE AND NON-
 CC MUSCLE CELLS IS NOT CLEAR.
 CC -!- SUBUNIT: DIMER OF AN ALPHA AND A BETA CHAIN.
 CC -!- DOMAIN: THE MOLECULE IS IN A COILED COIL STRUCTURE. THE SEQUENCE
 CC EXHIBITS A PROMINENT SEVEN-RESIDUES PERIODICITY.
 CC -!- TISSUE SPECIFICITY: BRAIN.
 CC -!- ALTERNATIVE PRODUCTS: THE TROPOMYOSIN ALPHA-TM GENE PRODUCES BY
 CC ALTERNATIVE SPlicing THE FOLLOWING PROTEINS: STRIATED MUSCLE TM,
 CC SMOOTH MUSCLE TM, BRAIN TMBR-1, TMBR-2, TMBR-3, FIBROBLAST TM-2,
 CC TM-3, TM-5A, AND TM-5B.
 DR EMBL; M34134; RNTMBR2A.
 DR PIR; B34787; B34787.
 DR PROSITE; PS00326; TROPOMYOSIN.
 DR BRAIN; COILED COIL; REPEAT; ALTERNATIVE SPlicing; MULTIGENE FAMILY.
 SQ SEQUENCE 251 AA; 28704 MW; 290473 CN;

Query Match 3.1%; Score 128; DB 7; Length 251;
 Best Local Similarity 23.5%; Pred. No. 1.75e-03;
 Matches 20; Conservative 28; Mismatches 34; Indels 3; Gaps 3;

Db 31 lqreid-gerklretaeadaslnrrlqlveeeldraderatalaqkleekaadeser 89
 Qy ::||: ; :::: ::||: ; ::||: ; ::||: ; ::||: ; ::||: ; ::||: ;
 318 LQRVIDSQAEKIKEL-DKEIRPFP-RQNWEADSKSSVLSQNRVTELESVDKSAGQVAR 375
 Db 90 gmkviesraqkdeekmeiqike 114
 Qy 376 NTGLLESQLSRHDQMLSVHDIRLAD 400

RESULT 14
 ID MTSP RAT STANDARD; PRT; 257 AA.
 AC P04452;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)

DE MYOSIN HEAVY CHAIN, PERINATAL SKELETAL MUSCLE (FRAGMENT).
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC ETHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N A.
 RM 85030490
 RA PERIASAMY M.; WIECZOREK D.F.; NADAL-GINARD B.;
 RL J. BIOL. CHEM. 259:1357-1358 (1984).
 CC -!- FUNCTION: MUSCLE CONTRACTION.
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM)
 CC AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATTER BE SPLIT FURTHER INTO
 CC 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).
 CC -!- DOMAIN: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -!- DEVELOPMENTAL STAGE: TRANSIENTLY EXPRESSED DURING PERINATAL
 CC SKELETAL MUSCLE DEVELOPMENT. IT IS FIRST DETECTED LATE IN FETAL
 CC LIFE, IS MAXIMALLY EXPRESSED AT THE END OF THE 1ST POSTNATAL
 CC WEEK, AND IS NOT FOUND IN THE ADULT. ALTHOUGH THEIR EXPRESSION
 CC TEMPORALLY OVERLAPS.
 CC -!- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,
 CC WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE, ARE
 CC CONSERVED.
 CC -!- THERE ARE 10 OR MORE MYOSIN HEAVY CHAIN GENES IN THE RAT.
 DR EMBL; K02111; RNMHC.
 DR PIR; A02911; A02991.
 KW MYOSIN; MUSCLE PROTEIN; COILED COIL; THICK FILAMENT; ACTIN-BINDING;
 KW ATP-BINDING; HEPTAD REPEAT PATTERN; MULTIGENE FAMILY.
 FT NON_TER 1
 FT DOMAIN <1 257 LIGHT MEROMYOSIN (LMM).
 SQ SEQUENCE 257 AA; 29843 MW; 288053 CN;

Query Match 3.1%; Score 131; DB 5; Length 257;
 Best Local Similarity 22.3%; Pred. No. 7.18e-04;
 Matches 39; Conservative 45; Mismatches 83; Indels 8; Gaps 8;

Db 42 qnasllna-kkkleldvsgiqseveeviqarnaaekakkaitdaammameikkcdtsa 100
 Qy ::||: ; ::||: ; ::||: ; ::||: ; ::||: ; ::||: ; ::||: ;
 257 OHVNLLKEWSNSLEKRVSLLONESWERKSISIHNQICSFETEIEROKEMIRNNESKL 316
 Db 101 hlermkknmeetvkqlhrld-eaqqlakggkq-ikklearvralleg-evenec-krn 156
 Qy 317 HLRVIDSQAEKIKEL-DKEIRPFP-RQNWEADSKSSVLSQNRVTELESVDKSAGQVAR 376
 Db 157 aeavkg-lrkherrelketylq-teedark-nvrlqdvlvdqkqavksykrgeea 208
 Qy 377 TGGLLESQLSRHDQMLSVHDIRLADMDLGQVLTASTNGVLWKRDKYKRRKQEA 431

RESULT 15
 ID TPM TRICO STANDARD; PRT; 284 AA.
 AC P15846;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)

Listing for Mary Haie

Tue Mar 19 12:36:57 1996

6

Listing for Mary Hale

Tue Mar 19 12:36:57 1996

4

CC will generate double deletion mutants. The invention is concerned with

CC the construction of attenuated viruses which have a reduced ability to

CC reactivate from latency. This can be achieved by functionally disabling

CC the expression of the EPO gene, or by disrupting the synthesis of the

CC LTR, or both. (see also 073500-1 and R60620-24)

SQ Sequence 69 AA;

Query Match 22.0%; Score 83; DB 11; Length 69;

CC Best Local Similarity 34.5%; Pred. No. 1.49e+00;

CC Matches 10; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

Db 24 cmhrfcipcmktwmqlnt-cplcnaklv 51

Oy 20 CGHRFCESCMALLSSSPKCTAQESTIV 48

RESULT 3
ID P31461 standard; peptide; 365 AA.

AC P31461;
DT 10-MAR-1993 (revised)
DE 09-JAN-1990 (first entry)

DE RPT-1 clone cDNA insert.

KW regulatory protein T lymphocyte-1; IL-2 receptor;

KW human lymphotrophic retroviruses; metal fingers.

OS Homo sapiens.

PN w08907652-A.
PD 24-AUG-1989.
PF 10-FEB-1989; US-154758.

PR PA (Dana-Farber Cancer Institute.

PI Cantor HI, Patarca R, Freeman GJ;
DR WPI; 89-263719736.

PT Regulatory Protein T lymphocyte-1 gene and corresp. protein - increase

PT gene expression of interleukin-2 receptor and of human lymphotrophic

PT retroviruses.

PS Claim 20; 58pp; English.

CC cDNA was prep. from T cells (fibroblast tumour) and 2PK3 (B cell

CC lymphoma). Host was pref. E. coli NRRL B-18297. Cysteine residues at

CC positions 15, 18, 35, 38, 55, 58, 115 and 118, and histidine residues at

CC positions 124 and 127 may be involved in metal finger formation.

CC Positions are given from the methionine (a.a 13 in the sequence given).

SQ Sequence 365 AA;

Query Match 21.8%; Score 82; DB 1; Length 365;
CC Best Local Similarity 34.6%; Pred. No. 1.82e+00;

CC Matches 9; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Db 27 cpcilelkkepvsadcnhsfcracit 52

Oy 5 CEKCHHVILCSKQTKCGRCESCMA 30

RESULT 4
ID R27535 standard; Protein; 560 AA.

AC R27535;
DT 09-MAR-1993 (first entry)

DE myl protein.

KW Retinoic acid receptor; RAR-alpha; myl; acute promyelocytic leukemia;

KW API; translocation; chromosome 17; chromosome 15; PCR; Primer;

KW Synthesis; Synthetic.

OS Synthetic.

PN W0921660-A.
PR (SIOK) SLOAN KETTERING INST CANCER.

PI Dmitrovsky E, Evans RM, Frankel S, Kazizuka A, Miller WH;
DR Warrell RP;
WPI; 92-349340/42.

DR N-PSDB; Q23134.

PT Marker for acute promyelocytic leukaemia and other neoplasias - comprising nucleic acid and encoded abnormal retinoic acid

KW receptor-alpha receptor

PT receptor-alpha receptor Disclosure; Page 48-51; 84pp; English.

CC The sequence given represents the amino acid sequence of one variation

CC of the myl protein. The gene encoding this protein is disrupted in a

CC translocation of a portion of the long arm of chromosome 17 onto the

CC long arm of chromosome 15 [t(15;17)(q21;q11-22)]. This causes a fusion

CC between Retinoic acid receptor (RAR)-alpha and myl which is

CC characteristic of acute promyelocytic leukaemia (APL). The breakpoint

CC region has been cloned and it has been shown that DNA rearrangements

CC are clustered in the region of the first intron of RAR-alpha. This

CC sequence was isolated by polymerase chain reaction (PCR). The primers

CC used for amplification of this sequence can also be used to amplify the

CC translocated region.

CC Sequence 560 AA;

SQ Query Match 20.2%; Score 76; DB 5; Length 560;

CC Best Local Similarity 34.5%; Pred. No. 5.90e+00;

CC Matches 10; Conservative 6; Mismatches 11; Indels 2; Gaps 2;

Db 56 rccqccgaaekc-pkllpc1htlcsgcglea 83

Oy 4 KECRCHL-VLCSPKQPCCGHRCESOMA 31

RESULT 5
ID R27533 standard; Protein; 797 AA.

AC R27533;
DT 09-MAR-1993 (first entry)

DE myl/RAR-alpha fusion.

KW Retinoic acid receptor; RAR-alpha; myl; acute promyelocytic leukemia;

KW API; translocation; chromosome 17; chromosome 15; PCR; Primer;

KW Synthesis; Synthetic.

OS Synthetic.

PN W0921660-A.
PR (SIOK) SLOAN KETTERING INST CANCER.

PI Dmitrovsky E, Evans RM, Frankel S, Kazizuka A, Miller WH;

DR Warrell RP;
WPI; 92-349340/42.

DR N-PSDB; Q23134.

PT Marker for acute promyelocytic leukaemia and other neoplasias - comprising nucleic acid and encoded abnormal retinoic acid

KW receptor-alpha receptor

Sizing 100 MHD 19 /Z-368/ 1998

Lisiting 101 (M&I) / 180 (M&I) / 13 / 2.3.0.3 / 1338

PS Disclosure; Fig 4; 84PP; English.
The sequence given shows a fusion between retinoic acid receptor
(RAR)-alpha and myl which is characteristic of acute promyelocytic
leukemia (APL). This is caused by a translocation of a portion of
the long arm of chromosome 17 onto the long arm of chromosome 15
[t(15;17)(q21;q11-22)]. The breakpoint region has been cloned and it
has been shown that DNA rearrangements are clustered in the region
of the first intron of RAR-alpha. This sequence was isolated by
polymerase chain reaction (PCR) using primers which correspond to
sequences both 5' and 3' to the breakpoint region.

Db	56	rcqgcqaekc-pklpcpltlcsgalea	83	Query Match 20.2%; Score 76; DB 5; Length 797;
Qy	4	KOEKCHL-VLGSPKQTECGHFRFCESCAA	31	Best Local Similarity 34.5%; Pred. No. 5.90e+00; Mismatches 6; Mismatches 11; Indels 2; Gaps 2;
RESULT	6			
ID	RJ0533	standard; Protein; 165 AA.		
AC	RJ0533;			
DT	12-APR-1991	(first entry)		
DE	Prod. of pmG4B12 used to isolate style-stigma specific gene STG4B12.			
KW	Hibrid vigour; pollination; Nicotiana tabacum "Petit Havana" SK1.			
OS	Nicotiana tabacum "Petit Havana" SK1.			
EP	EP-412006-A.			
PD	06-FEB-1991;			
PR	31-JUL-1990; 402196.			
PR	04-AUG-1989; EP-402224.			
PR	31-JUL-1990; EP-402196.			
PA	(PLAN-) PLANT GENETIC SYST.			
PI	De Greef W, Van Emmelo J, De Oliveira DE, De Souza MH;			
PI	Van Montagu M;			
DR	WP1; 91-038878/06.			
DR	N-PSDB; 010796.			
PT	Modified plant with transformed nuclear genome - obt'd. by using foreign DNA encoding prod. which disrupts metabolism, functioning and/or development of plant.			
PS	Diclaire, Fig 2B; 42pp; English.			
CC	The clone was produced by subcloning a style-stigma specific cDNA, 4B12, obt'd. from Prof. Goldberg of UCLA, into pGEM1, to produce pmG4B12. The clone was found to be stigma-style specific in Northern analysis. A probe from the clone was used to isolate the corresp. genomic sequence which is specifically expressed in style-stigma tissues of the female organ of tobacco. The correponding clone, pSTG4B12, contains STM4B12, an "STMG-type" gene. This and other similar genes can be used to produce new hybrid plants or seeds having a combination of desirable traits, and showing hybrid vigour. A female-sterile, male-fertile plant can be produced which favours cross-pollination.			
CC	See also RJ0531-35.			
SQ	Sequence 165 AA;			
Query Match	19.6%	Score 74; DB 2; Length 165;		
Best Local Similarity	47.8%	Pred. No. 8.69e+00; Mismatches 7; Indels 2; Gaps 2;		
Matches	11;	Conservative 3;		

RESULT	ID	Db	143	ncgcsnv-cspgqk-cafqfd	163
	AC	AC	:		:
	DT	DT	15-MAY-1995	(first entry)	
Pseudorabies virus early polypeptide 0 homologous protein domain.	R60622;	R60622;			
Pseudorabies virus; PRV; LIV; Large latency transcript; attenuated virus; vaccine; early protein 0; ED0; HSV-1 ICP0; protecting animals; deletion mutants; swine; cysteine-rich; zinc finger motif.	KW	KW			
Pseudorabies virus.	OS	OS			
Location/Qualifiers	FH	FH			
Misc difference 6	FT	FT			
/note= "cysteine residue that is part of the zinc finger motif"	FT	FT			
Misc difference 9	FT	FT			
/note= "cysteine residue that is part of the zinc finger motif"	FT	FT			
Misc difference 27	FT	FT			
/note= "cysteine residue that is part of the zinc finger motif"	FT	FT			
Misc difference 30	FT	FT			
/note= "cysteine residue that is part of the zinc finger motif"	FT	FT			
Misc difference 41	FT	FT			
/note= "cysteine residue that is part of the zinc finger motif"	FT	FT			
Misc difference 44	FT	FT			
/note= "cysteine residue that is part of the zinc finger motif"	FT	FT			
USDA) US SEC OF AGRIC.	PN	PN			
04-OCT-1994.	PD	PD			
11-SEP-1992; 945283.	PR	PR			
11-SEP-1992; US 945283.	PR	PR			
Cheung AK; Wesley RD; WPL; 94-31618/739.	PT	PT			
New Pseudorabies virus mutants for use in vaccine - having a deletion and/or insertion in the early protein 0 gene or large latency transcript gene	PT	PT			
Disclosure; Column 45:66; 43pp; English.	PS	PS			
R60622-24 show cysteine-rich zinc finger domains that is conserved among herpesviruses. This is the most important domain of the EPO gene. The DNA encoding this cysteine-rich domain (see Q73501) also encodes an amino acid sequence specific to the LIT gene in the opposite orientation. Thus deletion in the EPO gene other than the first 200 bp of EPO will automatically delete some of the DNA sequences encoding the LIT. EPO is essential for replication. LIT (large latency transcript) is the only gene expressed during PRV latency, and the IE180 gene is absolutely necessary for PRV replication. However there are 2 copies of IE180 in the genome. It is expected that PRV lacking one of the IE180 copies is viable. Deletions in the non-overlapping regions of these 3 genes will generate single deletion mutants, while deletions in overlapping regions will generate double deletion mutants. The invention is concerned with the construction of attenuated viruses which have a reduced ability to reactivate from latency. This can be achieved by functionally disabling	CC	CC			

CC	the expression of the EPO gene, or by disrupting the synthesis of the
CC	LTR, or both. (See also Q73500-1 and R60620-24)
SQ	Sequence 67 AA;
Query Match	19.4%; Score 73; DB 11; Length 67;
Best Local Similarity	36.2%; Pred. No. 1.0e-01;
Matches	17; Conservative 7; Mismatches 20; Indels 3; Gaps 2;
Match	17; Conservative 7; Mismatches 20; Indels 3; Gaps 2;
Db	6 cpcidlvaaatseqtlpmkbfcldcigrwltstacplckarvtsil 52
Qy	5 CERKHLVLCSPKQT-BGHRCESCMA--ALLSSSPKCTACQESIV 48
RESULT	8
ID	R60621; standard; Protein; 410 AA.
AC	R60621;
DT	15-MAY-1995 (first entry)
DE	Pseudorabies virus early polypeptide 0.
DE	KW Pseudorabies virus; PRV; LIV; large latency transcript;
DE	attenuated virus; vaccine; early protein 0; EPO; HSV-1 ICP0;
KW	protecting animals; deletion mutants; swine.
OS	Pseudorabies virus.
Key	Location/Qualifiers
FH	Domain 40..100
FT	/note= "cysteine rich zinc finger domain"
FT	Domain 46..49
FT	/label= zinc finger motif
FT	/note= "cysteine rich"
FT	Domain 67..70
FT	/label= zinc finger motif
FT	/note= "cysteine rich"
FT	Domain 81..84
FT	/label= zinc finger motif
FT	/note= "cysteine rich"
DR	US3352396-A.
PD	04-OCT-1994.
PF	11-SEP-1992; 945283.
PR	11-SEP-1992; US-945283.
PS	(USDA) US SEC OF AGRIC.
PI	Cheung AK; Wesley RD;
PI	WPI; 99-316187/39.
DR	N-PSDB; Q73501.
PT	New Pseudorabies virus mutants for use in vaccine - having a
PT	deletion and/or insertion in the early protein 0 gene or large
PT	latency transcript gene
Disclosure; Column 39-44; 43pp; English.	
CC	R60621 shows the early polypeptide (EPO) encoded by Q73501. The
CC	most important of the EPO gene is likely to be the cysteine-rich zinc
CC	finger domain from amino acids 40 to 100, since this region is conserved
CC	among other herpesviruses (eg. see R60622-24). The DNA encoding this
CC	cysteine-rich domain also encodes an amino acid sequence specific to the
CC	LTR gene in the opposite orientation. Thus deletion in the EPO gene
CC	other than the first 200 bp of EPO will automatically delete some of the
CC	DNA sequences encoding the LTR. EPO is nonessential for replication, LTR
CC	(large latency transcript) is the only gene expressed during PRV latency
CC	and the IE180 gene is absolutely necessary for PRV replication. However,
CC	there are 2 copies of IE180 in the genome. It is expected that PRV
CC	lacking one of the IE180 copies is viable. Deletions in the
CC	non-overlapping regions of these 3 genes will generate single deletion
CC	recombinants, while deletions in overlapping regions will generate double

CC	deletion mutants. The invention is concerned with the construction of attenuated viruses which have a reduced ability to reactivate from latency. This can be achieved by functionally disabling the expression of the EPO gene, or by disrupting the synthesis of the LIL, or both. (See also Q73500-1 and R0620-24)
CC	Sequence 410 AA;
CC	Query Match 19.4%; Score 73; DB 11; Length 410;
CC	Best Local Similarity 36.2%; Pred. No. 1.0⁺⁰¹;
CC	Matches 17; Conservative 7; Mismatches 20; Indels 3; Gaps 2;
Db	46 cpicidaataeqtpcmkfcldicqqrwlstacpolcnarvtsil 92
Qy	5 CICKHLVLCSPKOT-ECGHRFCESOMA--ALLSSSPKCTAACQESIV 48
RESULT	9
ID	R66033 standard; Protein: 574 AA.
AC	R66033;
DT	05-JUN-1995 (first entry)
DE	Human ARD 1.
KW	ADP-Ribosylation factor; ARF domain; ARD 1.
OS	Homo sapiens.
PN	W09424283-A.
DR	27-OCT-1994.
PF	15-APR-1994; US-044190.
PR	16-APR-1993; US-049252.
PA	19-APR-1993; US-049473.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
Mishina K, Moss J, Nightingale M, Tsuchiya M; MPI; 94-341862/42.	
NP	N-PSSDB; Q79326.
PT	GTP-binding protein ARD1 with ADP-ribosylation factor domain - useful as biochemical and diagnostic reagent
PS	Claim 5; Page 24-27; 52pp; English.
CC	The novel ARD 1 protein includes an 18 kDa region that exhibits significant homology to known ADP-ribosylation factors (ARFs) but lacks a 15 AA domain previously thought necessary for ARF stimulation. The 18 kDa region is called the functional domain and it is at the carboxy terminus. It is called ARD 1 for ARF domain. A recombinant truncated species contg. only the ARF domain following expression activated cholera toxin ADP-ribosyltransferase in a GTP-dependent manner, consistent with the conclusion that 15 AA adjacent to the amino terminus of ARF proteins are not required for toxin activation. cDNA was isolated from a human HL-60 lambda library screened with ARF 2B cDNA and a mixture of oligos denoted XARFC. Clone no. 76 was positive with the ARF 2B cDNA and negative with oligos specific for ARFs 1-6. The insert was sequenced. The insert (bps 606-2365 Q79326) included an ORF (1207-1722) encoding an ARF domain of 172 AA. Oligo JIR was used to screen a human fetal brain cDNA lambda ZAP library. Clones were found contg. nucleotides 7-1826 and 726-3225 of Q79326. In this sequence, about 1200 nts preceded the ARF region without a stop codon in the same ORF. To further characterize the 5' terminus of this cDNA, 5'-RACE was carried out with the poly A+ RNA from IMR-32 human neuroblastoma cells.
SQ	Sequence 574 AA;
Query Match 19.1%; Score 72; DB 12; Length 574;	
Best Local Similarity 33.3%; Pred. No. 1.28e+01;	

		Query Match	Score 69;	DB 12;	Length 146;
		Best Local Similarity	34.6%;	Pred. No. 2.26e+01;	
		Matches	9;	Mismatches	6;
Db	97	YKCSSG-NVITCDDKNDCESTCN-C	1 :: : : :	1 :: : : :	120
Qy	3	YKECKHIVLVISSPKTECGHRPCEC	28	: : : :	
RESULT	13				
ID	R53530	standard; protein;	185 AA.		
AC	R53530;				
DT	03-FEB-1995	(first entry)			
DE	Partial Thyroid hormone receptor-interacting protein - JIL2;	nuclear thyroid hormone interacting protein; TR; JIL1; JIL2;			
KW	LIM domain; consensus motif; lin-11; isl-1; mec-3;				
KW	transcriptional coactivator; treatment; diagnosis;				
KW	thyroid related disorders; modulation; thyroid hormone receptor;				
KW	nuclear hormone receptor; fusion proteins; isolation.				
OS	Homo sapiens.				
FH		Location/Qualifiers			
FT	Domain	30..81			
FT	/label= LIM_domain_I				
FT	Domain	90..151			
FT	/label= LIM_domain_II				
FT	Misc_difference	30			
FT	/note= "consensus Cys"				
FT	Misc_difference	33			
FT	/note= "consensus Cys"				
FT	Misc_difference	50			
FT	/note= "consensus His"				
FT	Misc_difference	53			
FT	/note= "consensus Cys"				
FT	Misc_difference	56			
FT	/note= "consensus Cys"				
FT	Misc_difference	77			
FT	/note= "consensus His"				
FT	Misc_difference	78			
FT	/note= "consensus Cys"				
FT	Misc_difference	81			
FT	/note= "consensus Asp"				
FT	Misc_difference	90			
FT	/note= "consensus Cys"				
FT	Misc_difference	117			
FT	/note= "consensus His"				
FT	Misc_difference	120			
FT	/note= "consensus Cys"				
FT	Misc_difference	123			
FT	/note= "consensus Cys"				
FT	Misc_difference	148			
FT	/note= "consensus Cys"				
FT	Misc_difference	151			
FT	/note= "consensus Cys"				
PN	W09410338-A.				
PD	11-MAY-1994.				
PF	29-OCT-1993; U10443.				
PR	30-OCT-1992; US-969136.				
PA	(GHRO) GEN HOSPITAL CORP.				
PI	Lee JW, Moore DD;				
DR	WPI, 94-199808/24.				

Qy 13 CSPKQTECHRFCECSMAALLSS 36

RESULT 15

ID R11108 standard; Protein; 555 AA.

AC R11108;

DT 23-MAY-1991 (first entry)

DE luciferase deduced from cDNA.

KW Vargula hilgendorfii.

OS Vargula hilgendorfii.

DR J03030678A.

PN 08-FEB-1991.

PF 29-JUN-1991, 167689.

PR 23-JUN-1989; JP-167689.

PA (OHSU-) OSAKA BIOSCIENCE K.

DR WPI, 91-08434312.

DR N-FSDB; Q10957.

PT DNA cdna coding luciferase derived from Vargula hilgendorfii - by

PT transfection of host cells and culturing.

PS Claim 1; Fig 3; 11pp; Japanese.

CC The cDNA encoding the protein is used to prepare vector PRSVL which is used to transfet COS cells (ATCC CRL1650) for expression of luciferase. The enzyme is useful for assays in biomedical or environmental fields.

CC Sequence 555 AA;

Query Match 18.3%; Score 69; DB 2; length 555; Best local Similarity 33.3%; Pred. No. 2.26e-01; Matches 8; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Db 39 ceakegeccidrcatckdilsg 62
| : | : || | : || :
Qy 13 CSPKQTECHRFCECSMAALLSS 36Search completed: Tue Mar 19 09:55:19 1996
Job time : 9 secs.

(TM)

Title: >US-08-404-832-3

Description: (1-49) from US08404832.pep

Perfect Score: 377

Sequence: 1 DKYKCKKHLVICSFSKQTECHRFCECSMAALLSSPKCTACQESIVK 49

Scoring table: PAM 150

Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

Database: pir46

pir46 ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4 8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1 13:unrev2

Statistics: Mean 30.545; Variance 66.324; scale 0.461

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	377	100.0	568	11 A55649	TINER-associated prot
2	120	31.8	377	11 A47380	RING finger-contain
3	109	28.9	1863	10 A54652	breast/ovarian cance
4	107	28.4	475	2 A37241	S2K autoantigen Ro/S
5	105	27.9	1042	11 S43379	RAG-1 protein - rabb
6	105	27.9	1042	11 S42511	RAG-1 protein - rabb
7	102	27.1	259	9 S23400	IAP53.3 protein - ye
8	102	27.1	487	3 DDBY18	DNA repair protein R
9	102	27.1	1043	11 A33754	recombination-activa
10	99	26.3	501	12 S24825	uve2 protein - Neuro
11	99	26.3	609	10 A43906	nuclear phosphoprote
12	98	26.0	1043	11 B33754	recombination-activa
13	97	25.7	477	9 S35494	C3HC4 type zinc fing
14	94	24.9	328	9 S00530	surface antigen CRP1
15	94	24.9	498	12 B42125	cysteine-rich surfac
16	94	24.9	624	10 S28418	probable zinc-bindin
17	94	24.9	1009	10 A56095	HIF16 protein - hum
18	94	24.9	1009	10 S49618	helicase-like transc
19	94	24.9	160	10 S26290	Rad-1 protein - chic
20	90	23.9	658	8 A44272	hypothetical protein
21	85	22.5	658	8 A44272	photomorphogenesis r
22	85	22.5	844866	10 R053.4 protein - Ca	R053.4 protein - Ca
23	85	22.5	1169	9 S31301	DNA repair protein R
24	83	22.0	315	6 D48560	immediate-early prot
25	83	22.0	598	7 S34203	heat shock protein 7
26	83	22.0	598	7 S28712	heat shock protein 7
27	83	22.0	630	10 A49656	estrogen-responsive
28	83	22.0	775	11 E0BE11	immediate-early prot
29	82	21.8	361	13 S17880	gene posterior sex c
30	82	21.8	365	11 A30891	regulatory protein r
31	82	21.8	1603	10 S17983	gene posterior sex c

Release 2.1D John F. Collins Biocomputing Research Unit.
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
 Distribution rights by Intelligent Genetics, Inc.

MSRCH_PP protein - protein database search, using Smith-Waterman algorithm
 Run on: Tue Mar 19 09:54:27 1996; Maspair time 3.63 Seconds
 Tabular output not generated.


```

##molecule_type DNA
##residues 1-1042 ##label FUS
##cross-references EMBL:N77666
SUMMARY
Query Match length 1042 #molecular-weight 118992 #checksum 8737
Best local Similarity 35.0%; Score 105; DB 11; Length 1042;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
QY 5 CERCHLVLICSPKOTEGGHRFCESMAALISSSSPKCTACQ 44

RESULT 7
ENTRY S53400 ##type complete
TITLE L8543.3 protein - yeast (Saccharomyces cerevisiae)
ORGANISM #formal name Saccharomyces cerevisiae
DATE 05-May-1995 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995

ACCESSIONS S53400
REFERENCE S53390
#authors Du, Z.
#submission submitted to the EMBL Data Library, February 1995
#description The sequence of S. cerevisiae cosmid 8543.
#deaccession S53400
#status preliminary
#molecule_type DNA
##residues 1-259 ##label DUZ

SUMMARY #length 259 #molecular-weight 29741 #checksum 8279
Query Match Best Local Similarity 27.1%; Score 102; DB 9; Length 259;
Matches 17; Conservative 6; Mismatches 18; Indels 1; Gaps 1;
Db 197 fktclckedkypspvurngyfgcscfaclmkqgt-kfifch 237
QY 3 YRCEKEHLVLICSPKOTEGGHRFCESMAALISSSSPKCTACQ 44

RESULT 8
ENTRY DBY18 ##type complete
TITLE DNA repair protein RAD18 - yeast (Saccharomyces cerevisiae)
ORGANISM #formal name Saccharomyces cerevisiae
DATE 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change
31-Mar-1995

ACCESSIONS S05002; S2263; S19481; JS0082
REFERENCE S05002
#authors Jones, J.S.; Weber, S.; Prakash, L.
#journal Nucleic Acids Res. (1988) 16:7119-7131
#title The Saccharomyces cerevisiae RAD18 gene encodes a protein
that contains potential zinc finger domains for nucleic
acid binding and a putative nucleotide binding sequence.
#cross-references MUID:88303333
#accession S05002
#molecule_type DNA
##residues 1-487 ##label JON
##cross-references EMBL:X12588

```

```

REFERENCE S22260
#authors Benit, P.; Chanet, R.; Fabre, F.; Faye, G.; Fukuhara, H.; Sor, F.; title Yeast (1992) 8:147-153 Sequence of the SUP61-RAD18 region on chromosome III of Saccharomyces cerevisiae.
#cross-references MUID:9221691
#accession S22263
#molecule_type DNA
##residues 1-487 ##label BEN
##cross-references EMBL:S93798
REFERENCE S19477
#authors Antoine, G.; Benit, P.; Chanet, R.; Fabre, R.; Faye, G.; Fukuhara, H.; Mathieu, A.; Sor, F. submitted to the Protein Sequence Database, March 1992
#accession S19481
#molecule_type DNA
##residues 1-487 ##label ANT
REFERENCE JS0082
#authors Chanet, R.; Magana-Schwencke, N.; Fabre, F.
#journal Gene (1988) 74:53-57
#title Potential DNA-binding domains in the RAD18 gene product of Saccharomyces cerevisiae.
#cross-references MUID:89232745
#accession JS0082
#molecule_type DNA
##residues 1-487 ##label CHA
GENETICS
#gene LISTA:RAD18
#map_position 3R
CLASSIFICATION #superfamily DNA repair protein RAD18
FEATURE ATP; DNA binding; DNA repair; zinc finger
KEYWORDS #region zinc finger motif\ATP; DNA binding; DNA repair; zinc finger
28-48 #region zinc finger motif\ATP; DNA binding; DNA repair; zinc finger
51-65 #region zinc finger motif\ATP; DNA binding; DNA repair; zinc finger
190-210 #region zinc finger motif\ATP; DNA binding; DNA repair; zinc finger
336 #binding site ATP (Lys) #status predicted
SUMMARY #length 487 #molecular-weight 55230 #checksum 1135
Query Match Best Local Similarity 27.1%; Score 102; DB 3; Length 487;
Matches 15; Conservative 6; Mismatches 18; Indels 1; Gaps 1;
Db 27 rchicidflkypvlpvcgfhfcslcarthlnq-pnprlc 65
QY 4 KCEKCHLVLICSPKOTEGGHRFCESMAALISSSSPKCTAC 43

RESULT 9
ENTRY A33754 ##type complete
TITLE recombinant-activating Protein 1 - human
ORGANISM #formal name Homo sapiens #common_name man
DATE 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
31-Dec-1993

ACCESSIONS A33754
REFERENCE A33754
#authors Schatz, D.G.; Oettinger, M.A.; Baltimore, D.
#journal Cell (1989) 59:1035-1048
#title The V(D)J recombination activating gene, RAG-1.

```

```

#cross-references MVID:9090604
#accession A3375g
##molecule_type mRNA
##residue 1-1043 #label SCA
##cross-references GB:M29414
#KEYWORDS DNA binding; zinc finger
#SUMMARY #length 1043 #molecular-weight 119115 #checksum 6389
#MMARY

Query Match 27.1%; Score 102; DB 11; Length 1043;
Best Local Similarity 32.5%; Pred. No. 4.2e-04;
Matches 13; Conservative 9; Mismatches 18; Indels 0; Gaps 0;
Db 141 eetcplclevlfkdpvnmvacqngnfcscidkaweqgssfacpecrei 188
Qy 1 DKYCKEKHLVLCSPKOTEGHRFCESMA-ALLSSSPKCTACOESI 47

RESULT 12
ENTRY B33754
TITLE recombination-activating protein 1 - mouse
ORGANISM #formal name Mus musculus #common name house mouse
DATE 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
31-Dec-1993
B33754

ACCESSIONS A33754
REFERENCE
#authors Schatz, D.G.; Oettinger, M.A.; Baltimore, D.
#JOURNAL Cell (1989) 59:1035-1048
#title The V(D)J recombination activating gene, RAG-1.
#cross-references MVID:9090604
#accession B33754

#molecule_type mRNA
##residues 1-1043 #label SCA
##cross-references GB:M29475
#length 1043 #molecular-weight 119490 #checksum 914
#KEYWORDS DNA binding; zinc finger
#SUMMARY #length 1043 #molecular-weight 119490 #checksum 914
#MMARY

Query Match 26.0%; Score 98; DB 11; Length 1043;
Best Local Similarity 32.5%; Pred. No. 1.6e-03;
Matches 13; Conservative 8; Mismatches 19; Indels 0; Gaps 0;
Db 293 cqicehlladpvtsckhlcircilrcikmgyspcr 332
Qy 3 YKCKEKHLVLCSPKOTEGHRFCESMA-ALLSSSPKCTACOES 46

RESULT 13
ENTRY S55494
TITLE C3HC4 type zinc finger protein - Emericella nidulans
ORGANISM #formal name Emericella nidulans, Aspergillus nidulans
DATE 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995
S55494

ACCESSIONS S55494
REFERENCE
#authors Iwanejko, L.A.; Cotton, C.M.; Jones, G.W.; Tomsatt, A.B.#
#title Strike, P.#
#cross-references MVID:92038424
#accession A43906
##molecule_type mRNA
##residues 1-609 #label RED

#MMARY

#cross-references MVID:9090604
#accession A43906
##molecule_type mRNA
##residues 1-609 #label RED

```

```

# residues 1-477 ##label IWA
##crosses-references EMBL:249834
SUMMARY
Query Match 25.7%; Score 97; DB 9; Length 477;
Best Local Similarity 34.9%; Pred. No. 2.24e-03;
Matches 15; Conservative 10; Mismatches 17; Indels 1; Gaps 1;
Match 15; Conservative 10; Mismatches 17; Indels 1; Gaps 1;
Db 29 reqcckdaffapvitescshfcscircrlsteg-keptcras 70
Qy 4 KCKCHLVLCSPKQTEGHRCESCMAALISSLSSPKCTACQES 46

RESULT 14
ENTRY S00530 #type fragment
TITLE surface antigen CRP170 - Giardia lamblia (fragment)
ORGANISM #formal name Giardia lamblia
DATE 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
18-Jun-1993
ACCESSIONS S00530
REFERENCE S00530
#authors Adam, R.D.; Aggarwal, A.; Lal, A.A.; de la Cruz, V.F.; McCutchan, T.J.; Nash, T.E.
#Journal J. Exp. Med. (1988) 167:105-118
#title Antigenic variation of a cysteine-rich protein in Giardia lamblia.
#cross-references MUID:88089405
#accession S00530
#molecule_type DNA
#residues 1-328 ##label ADA
##cross-references EMBL:X06741
#length 328 #checksum 3305
SUMMARY
Query Match 24.9%; Score 94; DB 9; Length 328;
Best Local Similarity 29.4%; Pred. No. 5.99e-03;
Matches 10; Conservative 10; Mismatches 13; Indels 1; Gaps 1;
Match 10; Conservative 10; Mismatches 13; Indels 1; Gaps 1;
Db 211 cpaatpgc-hnsdcgctenantnqadkctgckeg 243
Qy 13 cspkqtecgfrcescmaalisslsspkctacqes 46

RESULT 15
ENTRY B42125 #type complete
TITLE cysteine-rich surface antigen 170, CRP170 - Giardia lamblia
ORGANISM #formal name Giardia lamblia
DATE 04-Mar-1993 #sequence_revision 04-Mar-1993 #text_change
04-Mar-1993
ACCESSIONS B42125
REFERENCE B42125
#authors Adam, R.D.; Yang, Y.M.; Nash, T.E.
#Journal Mol. Cell. Biol. (1992) 12:1194-1201
#title The cysteine-rich protein gene family of Giardia lamblia: loss of the CRP170 gene in an antigenic variant.
#cross-references MUID:92166850
#content trophozoites
#accession B42125
#status preliminary
##molecule_type DNA
#residues 1-498 ##label ADA

```

```

*****cross-references NCIN:88424; NCBI:88431
*****note Sequence extracted from NCBI backbone
*****note #length 498 #molecular-weight 50111 #checksum 1456
*****note #summary

Query Match 24.9%; Score 94; DB 12; Length 498;
Best Local Similarity 29.4%; Pred. No. 5.99e-03;
Matches 10; Conservative 10; Mismatches 13; Indels 1; Gaps 1;
Db 96 cpaatpgc-hssedgctanmtqadkotgreq 128
Qy 13 CSPKTECCHRCSMMAILSSSPKCTACOQS 46

Search completed: Tue Mar 19 09:54:52 1996
Job time : 25 secs.

*****
***** Release 2.1D John F. Collins, Biocomputing Research Unit.
***** Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
***** Distribution rights by IntelliGenetics, Inc. (TM)
***** protein - protein database search, using Smith-Waterman algorithm
***** Run on: Tue Mar 19 09:54:02 1996; MacPar time 2.12 Seconds
***** 355.211 Million cell updates/sec

Tabular output not generated.

Title: >US-08-404-832-3
Description: (1-49) from US08404832.pep
Perfect Score: 377
Sequence: 1 DKYKCEKCHIVLCSPKQTECGHRCESCMMAILSSSPKCTACOQSIVK 49
Scoring table: PAM 150
Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%
listing first 45 summaries

Database:
swiss-prot31
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8

```

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description	Pred. No.
1	176	46.7	501	7	TRF2_MOUSE TNF RECEPTOR ASSOCIAT RING1 PROTEIN, CANCER TYPE 1	1.71e-19
2	120	31.8	377	6	RIN1_HUMAN RING1 PROTEIN, BREAST CANCER TYPE 1	1.11e-08
3	109	28.9	1863	1	BRCA1_HUMAN RING1 PROTEIN, BREAST CANCER TYPE 1	1.06e-06
4	107	28.4	475	6	RO52_HUMAN 52 KD RO PROTEIN (SJO 2.38e-06)	1.07e-06
5	105	27.9	1042	6	RAG1_RABBIT RAG1 (D) PROTEIN (RECOMBINATION A)	5.32e-06
6	102	27.1	487	6	RA18_YEAST DNA REPAIR PROTEIN RA18	1.76e-05
7	102	27.1	1043	6	RAG1_HUMAN RAG1 (D) PROTEIN (RECOMBINATION A)	1.76e-05
8	99	26.1	501	7	TVS2_NEUCR TVS2-2 PROTEIN	2.65e-05
9	98	26.0	1040	6	RAG1_MOUSE RAG1 (D) PROTEIN (RECOMBINATION A)	8.52e-05
10	94	24.9	328	1	C170_GIALLA SURFACE ANTIGEN CRP17	4.01e-04
11	94	24.9	624	1	A33_PLEWA ZINC-BINDING PROTEIN	4.01e-04
12	94	24.9	1041	6	RAG1_CHICK RAG1 (D) PROTEIN (RECOMBINATION A)	4.01e-04
13	90	23.9	160	7	YNN1_CAEEL HYPOOTHEICAL 18.7 KD	1.84e-03
14	85	22.5	834	8	YNC4_CAEEL HYPOOTHEICAL 97.1 KD	1.18e-02
15	85	22.5	1169	6	RADS_YEAST DNA REPAIR PROTEIN RA	1.18e-02
16	83	22.0	598	4	HST7_SBYV HEAT SHOCK PROTEIN 70	2.43e-02
17	83	22.0	775	4	ICPO_TSIV1 TRANSLATING TRANSCRIPT	2.43e-02
18	82	21.8	353	6	RPT1_MOUSE DOWN REGULATORY PROTEIN	3.49e-02
19	82	21.8	1603	6	PSC_DROME POSTERIOR SEX COMES P	3.49e-02
20	81	21.5	713	7	TS41_GIALLA MAJOR SURFACE-LABELED CISTERNE-RICH PROTEIN	5.00e-02
21	79	21.0	208	2	CRP2_RAT CRP2, RAT DP87 PROTEIN.	1.02e-01
22	79	21.0	555	2	ITB1_XENLA INTEGRIN BETA-1* SUBUNIT	1.44e-01
23	78	20.7	798	4	ITBO_XENLA INTEGRIN BETA-1* SUBUNIT	1.44e-01
24	78	20.7	324	1	EM11_MOUSE DNA-BINDING PROTEIN B	2.05e-01
25	77	20.4	324	1	EM11_MOUSE DNA-BINDING PROTEIN B	2.05e-01
26	77	20.4	326	1	EM11_MOUSE DNA-BINDING PROTEIN B	2.05e-01
27	76	20.2	274	8	YEV6_YEAST HYPOOTHEICAL 30.8 KD	2.99e-01
28	76	20.2	560	5	PM1B_HUMAN PUTATIVE TRANSCRIPT	2.99e-01
29	76	20.2	589	8	Y041_CAEEL HYPOOTHEICAL 68.7 KD	2.99e-01
30	76	20.2	611	5	PM12_HUMAN PROBABLE TRANSCRIPT	2.99e-01
31	76	20.2	633	5	PMIX_HUMAN PROBABLE TRANSCRIPT	2.99e-01
32	76	20.2	824	5	PM13_HUMAN PROBABLE TRANSCRIPT	2.99e-01
33	76	20.2	882	5	PM1L_HUMAN PROBABLE TRANSCRIPT	2.99e-01
34	76	20.2	1429	4	L112_CAEEL LIN-12 PROTEIN PRECUR	2.89e-01
35	75	19.9	115	8	YBR2_YEAST HYPOOTHEICAL 13.2 KD	4.08e-01
36	75	19.9	513	6	PSV_DCCD1 PRESPORSE VESICLE PROT	4.08e-01
37	74	19.6	803	4	ITB1_CHICK INTEGRIN BETA-1 PRECU	5.73e-01
38	74	19.6	1801	4	LMB2_RAT LAMININ BETA-2 CHAIN	5.73e-01
39	73	19.4	138	5	PA2B_VIPAA PHOSPHOLIPASE A2 PREC	8.03e-01
40	73	19.4	138	5	PA2C_VIPAA PHOSPHOLIPASE A2 PREC	8.03e-01
41	73	19.4	138	5	PA2A_VIPAA PHOSPHOLIPASE A2 PREC	8.03e-01
42	73	19.4	236	3	GLEM_TYME GIANT EXTRACELLULAR H	8.03e-01
43	73	19.4	342	4	ME18_MOUSE DNA-BINDING PROTEIN M	8.03e-01
44	73	19.4	344	4	ME18_HUMAN DNA-BINDING PROTEIN M	8.03e-01
45	73	19.4	410	4	ICFO_FNVIF TRANS-ACTING TRANSCRI	8.03e-01

ALIGNMENTS

RESULT 1
ID TRF2_MOUSE STANDARD; PRT; 501 AA.
AC P39429;
DT 01-FEB-1995 (REL. 31, CREATED)

Query Match 31.8%; Score 120; DB 6; Length 377;
 Best Local Similarity 34.8%; Pred. No. 1.1e-08; Indels 1; Gaps 1;
 Matches 16; Conservative 11; Mismatches 18; Index 1; Gaps 1;

Db 18 mcpicldmlkmttkeclhrfcadivvalrgnkeptccklv 63
 Qy :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

CC RESPONSIBLE FOR 45% OF INHERITED BREAST AND OVARIAN CANCER (BOC).
 CC OF INHERITED BREAST AND OVARIAN CANCER (BOC).
 CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
 DR EMBL; U146680; HS14660.
 DR MINI; 113705; 11TH EDITION.
 KW POLIMORPHISM; ANTI-ONCOGENE.
 KW ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN; DISEASE MUTATION;
 FT ZN FING 24 64 C3HC4-TYPE.
 FT VARIANT 64 64 C -> G (IN BREAST CANCER).
 FT VARIANT 356 356 C -> R.
 FT VARIANT 772 772 V -> A.
 FT VARIANT 1038 1038 E -> G.
 FT VARIANT 1040 1040 S -> N.
 FT VARIANT 1183 1183 K -> R.
 FT VARIANT 1443 1443 R -> G (IN BREAST CANCER).
 FT VARIANT 1613 1613 S -> G.
 FT VARIANT 1637 1637 P -> L (IN OVARIAN CANCER).
 FT VARIANT 1708 1708 A -> E (IN BREAST CANCER).
 FT VARIANT 1775 1775 M -> R (IN BREAST CANCER).
 SQ 1863 AA; 207720 MW; 16213725 CN;
 Query Match 28.9%; Score 109; DB 1; Length 1863;
 Best Local Similarity 39.6%; Pred. No. 1.0e-06; Indels 4; Gaps 4;
 Matches 19; Conservative 3; Mismatches 22; Index 1; Gaps 4;
 Db 24 cpicldmlkmttkeclhrfcadivvalrgnkeptccklv 70
 Qy 5 CERKHLVLCSPQTEGHRFCESMMAAIS-SSSP-KCTACESTIV 49
 RESULT 4
 ID R052 HUMAN STANDARD; PRT; 475 AA.
 AC P19474;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 20, LAST ANNOTATION UPDATE)
 DE 52 KD RO PROTEIN (Sjogren Syndrome Type A Antigen (SS-A)).
 GN SSA OR R052.
 OS HOMO SAPIENS (HUMAN).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC SCIENCE 265:120-122(1994).
 RN [1]
 RP VARIANTS BC GLY-64 AND GLY-143, AND VARIANTS ALA-772 AND ASN-1040.
 RA CASTILLA L.H., COYNE J.J., ERDOS M.R., HOSKINS K.F., CALZONE K.,
 RA GARNER J.E., BOYD J., LOUBIN M.B., DESHANO M.L., BRODY L.C.,
 RA COLLINS F.S., WEBER B.L.;
 RL NATURE GENET. 8:387-391(1994).
 RN [1]
 RP VARIANTS BC G-61, AND VARIANTS R-356; G-1038; N-1040; R-1183 & G-1613.
 RA FRIEDMAN L.S., OSTERMEYER E.A., SZABO C.I., DOWD P., LYNCH E.D.,
 RA RONKEL S.E., KING M.-C.;
 RL NATURE GENET. 8:399-404 (1994).
 CC -1- FUNCTION: NOT KNOWN. MAY REGULATE GENE EXPRESSION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- DISEASE: BREAST CANCER (BC) IS AN EXTREMELY COMMON MALIGNANCY,
 CC AFFECTING ONE IN EIGHT WOMEN DURING THEIR LIFETIME. A POSITIVE
 FAMILY HISTORY HAS BEEN IDENTIFIED AS MAJOR CONTRIBUTOR TO RISK OF
 DEVELOPMENT OF THE DISEASE, AND THIS LINK IS STRIKING FOR EARLY-
 CC ONSET BREAST CANCER. MUTATIONS IN BRCA1 ARE THOUGHT TO BE
 CC IDENTIFIED. THE COMMON FEATURE OF THESE PROTEINS IS THEIR
 CC ABILITY TO BIND RNA.

Tue Mar 19 12:36:58 1996

CC	-!- DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS OR PRIMARY SJOGREN'S SYNDROME OFTEN CONTAIN ANTIBODIES THAT REACT WITH NORMAL CELLULAR PROTEIN AS IF THIS ANTIGEN WAS FOREIGN.
CC	-!- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR.
CC	-!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
EMBL; M34551; HSROSSA.	DR
M62800; HSSARO.	DR
PIR; A37241; A37241.	DR
MINI; 105092; 11TH EDITION.	DR
PSD051B; ZINC FINGER C3HC4.	PROSITE
SYSTEMIC LUPUS ERYTHEMATOSUS; ZINC-FINGER; ANTIGEN; RNA-BINDING;	RIBONUCLEOPROTEIN; DNA-BINDING.
HYDROPHOBIC.	KW DOMAIN 1; 24
C3HC4-TYPE.	FT ZN FING 16 54
B BOX.	FT DOMAIN 92 123
LEUCINE-ZIPPER.	FT DOMAIN 211 232
HYDROPHOBIC.	FT DOMAIN 401 475
P -> A (IN REF. 2).	FT CONFLICT 52 52
PWYCN; 1162917 CN;	SEQUENCE 475 AA; 54169 MW;
Query Match 28.4%; Score 107; DB 6; Length 475;	Best Local Similarity 27.3%; Pred. No. 2, 38e-06;
Matches 12; Conservative 13; Mismatches 19; Indels 0; Gaps 0;	Matches 12; Conservative 13; Mismatches 19; Indels 0; Gaps 0;
Db 16 cpicldpfvepysieqgsfqeqeqaqvgqkqggsvpcvqrqlf 59	Db 16 cpicldpfvepysieqgsfqeqeqaqvgqkqggsvpcvqrqlf 59
5 CEKHLVLCSPKQTECGFECSCMAALLSSSPKCTACOESIV 48	5 CEKHLVLCSPKQTECGFECSCMAALLSSSPKCTACOESIV 48
RESULT 5	ID RAG1 RABBIT STANDARD; PRT; 1042 AA.
P34088; 01-FEB-1994 (REL. 28, CREATED)	ID DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DE V(D)J RECOMBINATION ACTIVATING PROTEIN.	DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
RAG-1.	DE V(D)J RECOMBINATION ACTIVATING PROTEIN.
ORCTOLAGUS CUNICULUS (RABBIT).	OS EURYORYTA; METAZOA; CHORDATA; TETRAPODA; MAMMALIA;
EUTHERIA; LAGOMORPHA.	OC
{1}	RN SEQUENCE FROM N.A.
RP 93354283	RP
FUSCHIOTTI P., HARINDRANATH N., MAGE R.G., MCCORMACK W.T.,	RA DHANARAJAN P., ROOK K.H.;
MOL. IMMUNOL. 30:1021-1032(1993).	RL
-!- FUNCTION: RAG1 & RAG2 SINERGISTICALLY ACTIVATE THE IMMUNOGLOBULIN V-D-J RECOMBINATION. V-D-J RECOMBINATION IS THE COMBINATORIAL PROCESS BY WHICH DEVELOPING LYMPHOCYTES BEGIN TO GENERATE THEIR ENORMOUS RANGE OF BINDING SPECIFICITIES FROM A LIMITED AMOUNT OF GENETIC INFORMATION.	CC
-!- SUBCELLULAR LOCATION: NUCLEAR.	CC
-!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.	CC
DR EMBL; M77666; OCRG1Q.	DR
PIR; S44379; S44379.	DR
PSD051B; ZINC FINGER C3HC4.	PROSITE
ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.	KW ZN FING 292 330 C3HC4-TYPE.
SEQUENCE 1042 AA; 119006 MW; 5538604 CN;	FT

ID	RAG1_HUMAN	STANDARD;	PRT;	1043 AA.
AC	P15918;			
DT	01-APR-1990	(REL. 14, CREATED)		
DT	01-APR-1990	(REL. 14, LAST SEQUENCE UPDATE)		
DT	01-NOV-1991	(REL. 20, LAST ANNOTATION UPDATE)		
GN	RAG1	RAG1 RECOMBINATION ACTIVATING PROTEIN.		
OS	HOMO_SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RM	90090604			
RA	SCHATZ D.G., OETTINGER M.A., BALTIMORE D.;			
RL	CELL 59:1035-1048(1989).			
CC	-!- FUNCTION: RAG1 & RAG2 SYNERGISTICALLY ACTIVATE THE IMMUNOGLOBULIN V-D-J RECOMBINATION. V-D-J RECOMBINATION IS THE COMBINATORIAL PROCESS BY WHICH DEVELOPING LYMPHOCYTES BEGIN TO GENERATE THEIR ENORMOUS RANGE OF BINDING SPECIFICITIES FROM A LIMITED AMOUNT OF GENETIC INFORMATION.			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-!- TISSUE SPECIFICITY: MATURING LYMPHOID CELLS.			
CC	-!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.			
CC	EMBL; M29474; HSRAIG1.			
DR	PIR; A33754; A33754.			
DR	HSSP; P28990; ICHC.			
DR	MIM; 179615; 11TH EDITION.			
DR	PROSITE; PS00518; ZINC FINGER C3HC4.			
KW	ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.			
FT	ZN FING 293 331 C3HC4-TYPE.			
SQ	SEQUENCE 1043 AA; 119115 MW; 5512008 CN;			
Query Match	Best Local Similarity 32.5%; Score 102; DB 6; Length 1043; Matches 13; Conservative 9; Mismatches 18; Indels 0; Gaps 0;			
Db	293 cqicebihiladpvertnckvlfvcrlckmgyspcr 332			
Qy	5 CEKCHLIVCSPKQTECGHRFCESCMALISSSSPKCTACQ 44			
RESULT	9			
ID	RAG1_MOUSE	STANDARD;	PRT;	1040 AA.
AC	P15919;			
DT	01-APR-1990	(REL. 14, CREATED)		
DT	01-APR-1990	(REL. 14, LAST SEQUENCE UPDATE)		
DT	01-FEB-1994	(REL. 28, LAST ANNOTATION UPDATE)		
DE	RAG-1			
DR	V(D)J RECOMBINATION ACTIVATING PROTEIN.			
GN	MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; RODENTIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RM	90090604			
RA	SCHATZ D.G., OETTINGER M.A., BALTIMORE D.;			
RL	CELL 59:1035-1048(1989).			
CC	[2]			
CC	MOTAGENESIS.			
CC	94111984			
RA	SADORSKY M.J., HESSE J.E., McBLANE J.F., GELLERT M.;			
RL	NUCLEIC ACIDS RES. 21:5694-5690(1993).			
CC	-!- FUNCTION: RAG1 & RAG2 SYNERGISTICALLY ACTIVATE THE IMMUNOGLOBULIN V-D-J RECOMBINATION. V-D-J RECOMBINATION IS THE COMBINATORIAL PROCESS BY WHICH DEVELOPING LYMPHOCYTES BEGIN TO GENERATE THEIR ENORMOUS RANGE OF BINDING SPECIFICITIES FROM A LIMITED AMOUNT OF GENETIC INFORMATION.			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-!- TISSUE SPECIFICITY: MATURING LYMPHOID CELLS.			
CC	-!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.			
DR	EMBL; M29475; MRAG1A.			
DR	PIR; B33754; B33754.			
DR	PROSITE; PS00518; ZINC FINGER C3HC4.			
KW	ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.			
FT	ZN FING 290 328 C3HC4-TYPE.			
SQ	SEQUENCE 1040 AA; 119160 MW; 5493298 CN;			
Query Match	Best Local Similarity 32.5%; Score 98; DB 6; Length 1040; Matches 13; Conservative 8; Mismatches 19; Indels 0; Gaps 0;			
Db	290 cqicebihiladpvertnckvlfvcrlckmgyspcr 329			

QY 5 CECKHVLICSPKTECGHRCESCMALLSSSPKCTACQ 44
 CC -!- DEVELOPMENTAL STAGE; IT FIRST APPEARS ON THE CHROMOSOME LOOPS
 CC AND IN THE NUCLEOPLASM OF THE GERMINAL VESICLE (GV). IT IS
 CC TRANSMITTED TO THE EGG AT GV BREAKDOWN AND APPEARS IN MANY BUT NOT
 CC ALL NUCLEI AT STILL LATER STAGES OF EMBRYOGENESIS.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
 DR EMBL; L04190; PWTWA32A;
 KW ZINC-FINGER; NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN; COILED COIL;
 KW HEPTAD REPEAT PATTERN; RNA-BINDING.
 FT DOMAIN 19 26 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 134 149 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT ZN FING 162 201 C3HC4-TYPE.
 FT DOMAIN 238 269 B BOX.
 FT DOMAIN 337 386 COILED COIL.
 FT DOMAIN 449 619 RFP-LIKE.
 FT SEQUENCE 624 AA; 71056 MW; 2025970 CN;
 SQ Query Match 24.9%; Score 94; DB 1; Length 624;
 Best Local Similarity 32.6%; Pred. No. 4.01e-04;
 Matches 15; Conservative 7; Mismatches 23; Indels 1; Gaps 1;
 Db 158 edltcplcrslfkepvilecqhnftkhcidkswessasafscpecke 203
 QY 1 DKYKECHVLICSPKTECGHRCESCMALLSSSP-KCTACQE 45
 RESULT 12
 ID RAG1 CHICK STANDARD; PRT; 1041 AA.
 AC P24271;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-OCT-1992 (REL. 21, LAST ANNOTATION UPDATE)
 DE V(D)J RECOMBINATION ACTIVATING PROTEIN.
 GN RAG-1.
 OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEognathae;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RM 91098648
 RA CARLSON L.M., OTTINGER M.A., SCHATZ D.G., MASTELLER E.L.,
 RA HURLEY E.A., MCCORMACK W.I., BALTIMORE D., THOMSON C.B.;
 RL CELL 64:2201-208(1991).
 CC -!- FUNCTION: RAG1 & RAG2 SYNERGISTICALLY ACTIVATE THE IMMUNOGLOBULIN
 CC V-D-J RECOMBINATION. V-D-J RECOMBINATION IS THE COMBINATORIAL
 CC PROCESS BY WHICH DEVELOPING LYMPHOCYTES BEGIN TO GENERATE THEIR
 CC ENORMOUS RANGE OF BINDING SPECIFICITIES FROM A LIMITED AMOUNT OF
 CC GENETIC INFORMATION.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
 DR EMBL; M58530; GSBRAGIG.
 DR PIR; S42509; S42509.
 DR PROSITE; PS00518; ZINC FINGER_C3HC4.
 KW ZINC-FINGER; DNA-BINDING; NUCLEAR PROTEIN.
 FT ZN FING 284 327 C3HC4-TYPE.
 SQ SEQUENCE 1041 AA; 119916 MW; 544207 CN;
 Query Match 24.9%; Score 94; DB 6; Length 1041;
 Best Local Similarity 30.8%; Pred. No. 4.01e-04;

RESULT 11
 ID A33 PLEWA STANDARD; PRT; 624 AA.
 AC 002084;
 DT 01-OCT-1994 (REL. 30, CREATED)
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE ZINC-BINDING PROTEIN_A33.
 OS PLAKOPOLELES WALMUT (IBERIAN RIBBED NEWT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA;
 OC CRUDATA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=OVARY;
 RM 93154311
 RA BELLINI M., LACROIX J.-C., GAIL J.G.;
 RL EMB0 J. 12:1107-114 (1993).
 CC -!- FUNCTION: MAY BE A NUCLEAR REGULATORY PROTEIN THAT IS STORED
 CC IN THE GERMLINAL VESICLE FOR USE DURING EARLY EMBRYOGENESIS
 CC AND MAY PLAY A ROLE IN THE SYNTHESIS OR PROCESSING OF PRE-mRNA
 CC DURING OOGENESIS.

Listing for May Halle
100 Mar 19 12:30:58 1998

CC MAINTENANCE OF WILD-TYPE RATES OF INSTABILITY OF SIMPLE
 REPETITIVE SEQUENCES SUCH AS POLY(GT) REPEATS.
 CC
 CC -|- SUBCELLULAR LOCATION: NUCLEAR.
 CC -|- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
 DR EMBL; M96644;
 DR EMBL; S43248; S43248.
 DR EMBL; S46103; S46103.
 DR PIR; S31301; S31301.
 DR PIR; S26983; S26983.
 DR LISTA; SC00524; RAD5.
 KW DNA DAMAGE; DNA REPAIR; ZINC-FINGER; NUCLEAR PROTEIN; ZINC-FINGER; DNA-BINDING;
 KW HELICASE; ATP-BINDING.
 FT DOMAIN 42 60 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 303 315 ARG/LYS-RICH (BASIC).
 FT NP BIND 532 539 ATP (POTENTIAL).
 FT SITE 681 684 DEGH BOX.
 FT ZN FING 914 960 C3HC4-TYPE.
 FT CONFLICT 478 478 Q -> R (IN REF. 2).
 FT CONFLICT 635 635 T -> N (IN REF. 2).
 FT CONFLICT 846 846 G -> S (IN REF. 2).
 FT CONFLICT 898 898 R -> S (IN REF. 2).
 FT CONFLICT 973 973 V -> A (IN REF. 2).
 FT CONFLICT 1063 1063 A -> R (IN REF. 2).
 SQ SEQUENCE 1169 AA; 134001 MW; 7092987 CN;

Query Match 22.5%; Score 85; DB 6; Length 1169;
 Best Local Similarity 75.0%; Pred. No. 1.18e-02;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db	930 tecghafceekl	941
Qy	18	TEOGHAFCECM 29

Search completed: Tue Mar 19 09:54:09 1996
 Job time : 7 secs.